

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 06:02:46 ; Search time 23.3657 Seconds
(without alignments)
1325.635 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600

Perfect score: 101

Sequence: 1 atattggcctggtgctcc.....tatgtaaagcttcaagtt 101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

al number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

- 1: /cn2.5/ptodata/1/ina/5A_COMB.seq:*
- 2: /cn2.5/ptodata/1/ina/5B_COMB.seq:*
- 3: /cn2.5/ptodata/1/ina/6A_COMB.seq:*
- 4: /cn2.5/ptodata/1/ina/6B_COMB.seq:*
- 5: /cn2.5/ptodata/1/ina/8B_COMB.seq:*
- 6: /cn2.5/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	32	31.7	300	3	US-09-009-494-3	Sequence 3, Appli
C 2	32	31.7	300	4	US-09-010-233-5	Sequence 5, Appli
C 3	32	31.7	300	4	US-09-010-232-1	Sequence 1, Appli
C 4	27.8	27.5	4597	4	US-08-961-527-175	Sequence 175, App
C 5	27.8	27.5	4797	4	US-09-419-568F-25	Sequence 25, Appl
C 6	27.8	27.5	4797	4	US-09-354-243B-25	Sequence 25, Appl
C 7	27.8	27.5	5643	1	US-08-144-602B-4	Sequence 4, Appli
C 8	27.4	27.1	3340	4	US-09-450-852-3	Sequence 3, Appli
C 9	26.8	26.5	1894	3	US-09-082-310-3	Sequence 3, Appli
C 10	26.8	26.5	1894	4	US-09-575-203-3	Sequence 3, Appli
C 11	26.4	26.1	3373	4	US-09-457-066-42	Sequence 42, Appl
C 12	26.4	26.1	3826	4	US-09-302-620B-90	Sequence 90, Appl
C 13	26.4	26.1	5599	2	US-08-477-451-9	Sequence 9, Appli
C 14	26.4	26.1	5599	2	US-08-477-451-13	Sequence 13, Appl
C 15	26.4	26.1	19332	2	US-08-477-451-25	Sequence 25, Appl
C 16	26.2	25.9	16995	4	US-08-961-527-82	Sequence 82, Appl
C 17	26	25.7	2343	4	US-09-307-143-5	Sequence 5, Appli
C 18	26	25.7	4673	1	US-07-638-431-1	Sequence 1, Appli
C 19	26	25.7	4673	5	PCT-US92-00018-1	Sequence 1, Appli
C 20	25.8	25.5	2821	4	US-09-484-970B-44	Sequence 44, Appl
C 21	25.4	25.1	1351	4	US-09-002-114-2	Sequence 2, Appli
C 22	25.4	25.1	2664	2	US-08-942-819-1	Sequence 1, Appli
C 23	25.4	25.1	11478	3	US-08-981-803-29	Sequence 29, Appl
C 24	25.4	25.1	11478	4	US-08-983-440-29	Sequence 29, Appl
C 25	25.2	25.0	361	4	US-09-018-584A-16	Sequence 16, Appl
C 26	25.2	25.0	1001	4	US-09-641-638-294	Sequence 294, App
C 27	25.2	25.0	1110	2	US-08-719-758-1	Sequence 1, Appli

Query Match

31.7%; Score 32; DB 3; Length 300;

ALIGNMENTS

RESULT 1

US-09-009-494-3/c

: Sequence 3, Application US/09009494

: Patent No. 6150349

: GENERAL INFORMATION:

: APPLICANT: Black, Michael T.

: APPLICANT: Lawlor, Elizabeth J.

: APPLICANT: Lewis, Ceri J.

: TITLE OF INVENTION: No. 6150340e1 Compounds

: NUMBER OF SEQUENCES: 6

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Dechert, Price & Rhoads

: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre

: CITY: Philadelphia

: STATE: PA

: COUNTRY: USA

: ZIP: 19103-2793

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FastSeq for Windows Version 2.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/009,494

: FILING DATE:

: CLASSIFICATION:

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 60/037,857

: FILING DATE: 07-FEB-1997

: APPLICATION NUMBER: 60/044,365

: FILING DATE: 28-APR-1997

: APPLICATION NUMBER: 60/044,366

: FILING DATE: 28-APR-1997

: ATTORNEY/AGENT INFORMATION:

: NAME: Falk, Stephen T

: REGISTRATION NUMBER: 36,795

: REFERENCE/DOCKET NUMBER: GM50028

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 215-994-2488

: TELEFAX: 215-994-2222

: TELEX:

: INFORMATION FOR SEQ ID NO: 3:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 300 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: double

: TOPOLOGY: linear

: US-09-009-494-3

Best Local Similarity	68.8%;	Pred. No.	0.19;		
Matches	44;	Conservative	0;	Mismatches	20;
				Indels	0;
				Gaps	0;

	Matches	44:	Conservative	0:	Mismatches	20:	Indels	0:	Gaps
QY	38	ATGCCITTCCTCTCTTGAATATCTTCTGTTCTTCATAGATATCTAAATGCTTTCA	97						
Db	104	ATGCTTTAAATGTTGCCCATTTCTCGTTCCTTCAGAGAGAAATTTGAGCTCTGCA	45						
QY	98	AGTIT	101						
Db	44	ACAT	41						

DECLASS

RESCLI 2
US-09-010-233-5/c
; Sequence 5, Application US/09010233
; Patent No. 6200774
; GENERAL INFORMATION:
; APPLICANT: Black, Michael T.
; APPLICANT: Lawlor, Elizabeth J.
; APPLICANT: Lewis, Carl J.
; TITLE OF INVENTION: No. 6200774e1 Compounds
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Ar
; CITY: Philadelphia
; STATE: PA

```

:
: COUNTRY: USA
:
: ZIP: 19103-2793
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Diskette
:
: COMPUTER: IBM Compatible
:
: OPERATING SYSTEM: DOS
:
: SOFTWARE: FASTSEQ for Windows version 2.0
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/09/010,233
:
:

```

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/037,857
 FILING DATE: 07-FEB-1997
 APPLICATION NUMBER: 60/044,365
 FILING DATE: 28-APR-1997
 APPLICATION NUMBER: 60/044,366
 FILING DATE: 28-APR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Falk, Stephen I
 REGISTRATION NUMBER: 35,795
 REFERENCE/DOCKET NUMBER: GM50027
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-994-2488
 TELEFAX: 215-994-2222
 TELEVISION:

1444A.
INFORMATION FOR SEC TO NO. 5.

```

; INFORMATION FOR SEQ ID NO: 3
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 300 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-02-010-233-5

```

Query Match 31.7%; Score 32; DB 4; Length 320;
Best Local Similarity 69.8%; Pred. No. 0.19;
Matches 44; Conservative 0; Mismatches 20; Indels.

		Matches	44	Conservative	0	Mismatches	20	Indels	0	Gaps
QY	38	ATGCTTTTCCCTCTCTTGATTAATCTTCTTCGTTCTTCATAGATAGTAAATGCTTCA	97							
Db	104	ATGCTTTTAAATGTTGCCAATTCCTCGGTTCTTCAGGAGAAATTTGAAGCTCTTCA	45							
QY	98	AGTT	101							
Db	44	AGAT	41							

Db 44 AGAT 41

RESULT 3
US-09-010-232-1/C
: Sequence 1, Application US/09010232
: Patent No. 6248557
: GENERAL INFORMATION:
: APPLICANT: Black, Michael T.
: APPLICANT: Lawlor, Elizabeth J.
: APPLICANT: Lewis, Carl J.

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dechert, Price & Rhoads
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103-2793
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: fastseq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/010.232

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/937,857
FILING DATE: 07-FEB-1997
APPLICATION NUMBER: 60/044,366
FILING DATE: 28-APR-1997
APPLICATION NUMBER: 60/044,365
FILING DATE: 28-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen I
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GMS0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEF:

```

; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 300 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;
; US-00-010-232-1

```

Query Match 31.7%; Score 32; DB 4; Length 300;
Best Local Similarity 68.8%; Pred. No. 0.19;
Matches 44; Conservative 0; Mismatches 20; Indels

QY: 38 ATGCGTTTTCGTCTTAATTTCTTGTTCTTCATTAGATATGTAAATGCTTTCA 97
||||| | ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 104 ATGCTTTTAACTGGTGGCCATTCTTCGGTTCTTCAGGAGAAATTGAAGTCTTCCA 45

QY 98 AGTT 10

Db 44 AGAT 41

RESULT 4
IIS-08-961-527-176

: Sequence 175, A

Patent No. 6420
GENERAL INFORMATION

APPLICANT:
 TITLE OF INV:

NUMBER OF SE
CORRESPONDEN

CORRESPONDENCE: ADDRESS:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 05:54:07 ; Search time 754.132 Seconds
(without alignments)
2086.058 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600
Perfect score: 101
Sequence: 1 atattggtcctgtgtctcc.....tatgtaaatgcttcacagtt 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
cal number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_esti:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estim:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pin:*
- 21: em_gss_vit:*
- 22: em_gss_fun:*
- 23: em_gss_nam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
C 1	101	100.0	432	9	AA421341	AA421341 zu6cd11.f
C 2	74.4	73.7	513	9	AI025332	AI025332 ov74g03.x
C 3	33.2	32.9	232	17	BH757424	BH757424 SALK_0561
C 4	33.2	32.9	508	17	AQ779483	AQ779483 HS_3001.A
C 5	33	32.7	476	17	A2809047	A2809047 2M072E19
C 6	32.2	31.9	1197	12	BG538053	BG538053 602563562

C 7	32	31.7	570	13	B1620394	B1620394 603036284
C 8	32	31.7	568	17	BH768691	BH768691 BMBAC361D
C 9	31.6	31.3	403	17	FR0023141	AL016000 F.rubripe
C 10	31.6	31.3	435	9	AL726759	AL726759 AL726759
C 11	31.6	31.3	555	17	AO651385	AO651385 Sheared D
C 12	31.6	31.3	578	9	AL514767	AL514767 AL514767
C 13	31.6	31.3	911	17	AO899744	AO899744 HS_5234..A
C 14	31.4	31.1	554	17	AQ531126	AQ531126 RECI-23-2
C 15	31.2	30.9	452	10	AW988030	AW988030 uf91601.Y
C 16	31.2	30.9	514	17	BH824692	BH824692 BACP20-L
C 17	31.2	30.9	535	17	A2636018	A2636018 1M0457009
C 18	31.2	30.9	692	17	BH834121	BH834121 BACP66-K1
C 19	31.2	30.9	728	17	BH826207	BH826207 BACP23-E
C 20	31.2	30.9	808	13	B1088928	B1088928 602854293
C 21	31.2	30.9	900	9	AL547213	AL547213 AL547213
C 22	31.2	30.9	1123	12	B3701181	B3701181 602680785
C 23	31	30.7	414	17	A2868761	A2868761 2M018C509
C 24	31	30.7	636	12	B528560	B528560 602043533
C 25	31	30.7	645	17	A2458131	A2458131 2M0222D02
C 26	31	30.7	881	17	BH154433	BH154433 ENTBP47TR
C 27	31	30.7	941	17	A2534286	A2534286 ENTBM45TF
C 28	31	30.7	1101	17	CNS01501	AL104844 Drosophila
C 29	30.8	30.5	235	9	AI772083	AI772083 EST253183
C 30	30.8	30.5	349	10	AW177332	AW177332 CM1-C1012
C 31	30.8	30.5	512	17	AQ765687	AQ765687 HS_5372..A
C 32	30.8	30.5	721	12	B3762662	B3762662 602734558
C 33	30.8	30.5	855	17	CNS012BX	AL101367 Drosophila
C 34	30.6	30.3	106	12	B5827565	B5827565 RC2-E1001
C 35	30.6	30.3	482	9	AI317159	AI317159 u135903.Y
C 36	30.6	30.3	534	17	AQ735325	AQ735325 HS_2011..A
C 37	30.6	30.3	583	17	AQ531230	AQ531230 RECI-11-3
C 38	30.6	30.3	596	17	BH255245	BH255245 CH230-123
C 39	30.6	30.3	713	17	A2631576	A2631576 1M0485M18
C 40	30.6	30.3	760	17	A2628465	A2628465 1M0486N05
C 41	30.6	30.3	1067	17	AG029559	AG029559 Pan trogl
C 42	30.4	30.1	309	10	AW177337	AW177337 CM1-C1012
C 43	30.4	30.1	320	10	AW177341	AW177341 CM1-C1012
C 44	30.4	30.1	320	10	AW177381	AW177381 CM4-CT012
C 45	30.4	30.1	321	10	AW177379	AW177379 CM4-CT012

ALIGNMENTS

RESULT 1
LOCUS AA421341/C 432 bp mRNA linear EST 09-NOV-1997
DEFINITION zu6cd11.f1 Soares_testis_NH1 Homo sapiens cDNA clone IMAGE:731061
5' similar to gb:X51602.cdsl VASCULAR ENDOTHELIAL GROWTH FACTOR
RECEPIOR 1 (HUMAN);, mRNA sequence.

ACCESSION AA421341
VERSION AA421341.1 GI:2100333
KEYWORDS EST.
SOURCE Human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 432)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gaisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tag,F., Theising,B.,
White,T., Wylie,T., Waterston,R. and Wilson,R.

TITLE

WashU-NCI human EST Project

JOURNAL

Unpublished (1997)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL : contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 801 Std Error: 0.00


```

BASE COUNT      73 a    25 c    40 g    79 t    15 others
ORIGIN

Query Match      32.9%; Score 33.2; DB 17; Length 232;
Best Local Similarity 61.6%; Pred. No. 1.7e-02;
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 10 CTGTGTCCTGACATCAGATGAGCTGCTTTCTGCTGCTGATGATGCTTCTTCT 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 6 CATGAACCTCTGAGATGCTGAGATGAGAAAGTTTGGCCATGTTTGCCATCTTTATTT 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 70 TTTCTCAATTAGATGATTAAGTCTTT 95
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 66 CACTTCGTGATAGTAAAGATTT 91
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
A0779483/C
LOCUS      508 bp    DNA    linear    GSS 02-AUG-1999
DEFINITION HS-3001 AL-B02-WTC CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate-3001 Col-3 Row-C, DNA sequence.
ACCESSION  A0779483
VERSION    A0779483.1 GI:5682443
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 508)
AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Kallier,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE   99380589
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end web server: http://www.htsc.washington.edu
            Plate: 3001 row: C column: 3
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 508.
            Location/Qualifiers
                1..508
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="Plate-3001 Col-3 Row-C"
                /clone_lib="CIT Approved Human Genomic Sperm Library D"
                /sex="male"
                /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                E-coli DH10B"
BASE COUNT      200 a    93 c    74 g    137 t    4 others
ORIGIN

Query Match      32.9%; Score 33.2; DB 17; Length 508;
Best Local Similarity 61.6%; Pred. No. 1.6e-02;
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 15 GTCTCTGACTCATATAGAGCTGAGTGGCTTTCTGCTGCTGATGATGCTTCTGCT 74
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 212 GTGCTCTCTTCAICAGTATGATTTAGTTTTCCTGTATAGATCTTCTTCTTATGII 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 75 CATTAGATATGTAATGCTTTCAAGT 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 152 CATTGATTCGTAGGTAGTTATAT 127
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 5
A2809047
LOCUS      476 bp    DNA    linear    GSS 20-FEB-2001
DEFINITION 2M0072E19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0072E19 R, DNA sequence.
ACCESSION  A2809047
VERSION    A2809047.1 GI:12975012
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 476)
REFERENCE  1
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., DuVal,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenan,E., Pedersen,I., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tirgey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0072 row: E column: 19
            Seq primer: CACACAGAAACAGCTATGACC
            Class: Plasmid ends
            High quality sequence stop: 476.
            Location/Qualifiers
                1..476
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGC2M0072E19"
                /clone_lib="Mouse 10kb plasmid UUGC1M library"
                /sex="Male"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, P-
                /note="Vector: FWD42nv; Purified genomic DNA from M.
                musculus C57BL/6J (male) was obtained from the Jackson
                Laboratory Mouse DNA Resource
                (http://www.jax.org/resources/documents/dnares/). The DNA
                was hydrodynamically sheared by repeated passage through a
                0.005 inch orifice at constant velocity. The sheared DNA
                was blunt end-repaired with T4 DNA polymerase and T4
                polynucleotide kinase. Adaptor oligonucleotides were
                ligated to the blunt ends in high molar excess. The
                adaptor DNA was purified and size-selected for a 9.5 to
                10.5 kb range using preparative agarose gel
                electrophoresis. Vector DNA was prepared from a derivative
                of pWD42 (gi14732114|gb|AF129072.1), a copy-number
                inducible derivative of plasmid R1. The vector was ligated
                with adaptors complementary to the insert adaptors and
                purified. The sheared, adaptor mouse DNA was annealed to
                adaptor vector DNA, and transformed into
                chemically-competent E. coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
BASE COUNT      111 a    123 c    100 g    142 t
ORIGIN

Query Match      32.7%; Score 33; DB 17; Length 476;
Best Local Similarity 63.0%; Pred. No. 1.8e-02;
Matches 51; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 17 CTCCTGACTCATTAGAGCTGGAGTCCCTTCTGCTGCTGATGATGCTTCTGCTTCTCA 76
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 152 CTCACGCTGGCAGAGCTGGTGACAGCCCTGCTGCTGCTTCTTCTTCTTCTCT 211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 Sequenced from the Brugia malayi BAC library constructed by Claire Whittom and Dr Mike Quail. The sequence was generated by The Pathogen Sequencing Unit, The Sanger Institute, Cambridge, UK in collaboration with Mark Blaxter, ICRP, University of Edinburgh, Edinburgh, UK.
 Seq primer: T7 (TATACGACTCACTATAGG)
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..668
 /organism="Brugia malayi"
 /strain="TRS"
 /db_xref="taxon:6279"
 /clone_lib="Brugia malayi Genomic BAC Library 3"
 /sex="Mixed (male and female)"
 /tissue_type="whole parasite"
 /dev_stage="microfilaria (L1)"
 /notes="Vector: pBACE3.6; Site 1: BamH I; Brugia malayi genomic DNA was partially cleaved with Sau3A I and size fractionated. 7,392 clones were generated with mean insert size ~48 kbp. The library was constructed by Claire Whittom, Blaxter Nematode Genetics Lab, University of Edinburgh, UK, and Dr Mike Quail, The Pathogen Sequencing Unit, The Sanger Centre, Cambridge, UK."
 235 a 98 c 122 g 213 t

BASE COUNT

ORIGIN

Query Match 31.7%; Score 32; DB 17; Length 668;
 Best Local Similarity 58.3%; Pred. No. 3e+02;
 Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 5 GTGGCTGTGCTCTCGTCACTATTAGAGCTGGATGCTTTCTCTGTTGTAATCTT 64
 DQ 1.3 GTGGCTGTGCTCTCGTCACTATTAGAGCTGGATGCTTTCTCTGTTGTAATCTT 64

QY 65 TGTGTTCTTCATTAGATATGTAATGCTTTCAAGT 100
 DQ 1.1 TGTGTTCTTCATTAGATATGTAATGCTTTCAAGT 100

QY 53 TTCTGTTTTCGAATGATATGACGTCTGCTCAAGT 18
 DQ 1.1 TTCTGTTTTCGAATGATATGACGTCTGCTCAAGT 18

RESULT

FR0023141/c

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

PRIMER: KS

DESCR:

One pass dye-terminator sequencing of cosmid cloned genomic sequence.

FEATURES

Location/Qualifiers

1..403

/organism="Takifugu rubripes"

/db_xref="taxon:31033"

/clone="089H1aC2"
 /clone_lib="cosmid 089H1a"
 143 a 76 c 81 g 102 t

Query Match 31.3%; Score 31.6; DB 17; Length 403;
 Best Local Similarity 62.8%; Pred. No. 4e+02;
 Matches 49; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 ATAAGTGGCTGGTCTCTGACTCATTAGAGCTGGATGCTTTCTCTGTTGTAAT 60
 DQ 1.1 ATAAGTGGCTGGTCTCTGACTCATTAGAGCTGGATGCTTTCTCTGTTGTAAT 60

QY 91 ATTCTGCTGGTCTCTGACTCATTAGAGCTGGATGCTTTCTCTGTTGTAAT 32
 DQ 1.1 ATTCTGCTGGTCTCTGACTCATTAGAGCTGGATGCTTTCTCTGTTGTAAT 32

QY 61 TCTTCTGTTCTTCAAT 78
 DQ 1.1 TCTTCTGTTCTTCAAT 78

QY 31 TTTAACCCTTCTAGCAT 14
 DQ 1.1 TTTAACCCTTCTAGCAT 14

RESULT 10

AL726759/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 Cyprinidae; Danio.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..435

/organism="Danio rerio"

/db_xref="taxon:7955"

/clone="BN0AA0882F04"

/tissue_type="inner ear"

/dev_stage="embryonic"

/note="subtracted cDNA library"

143 a 102 c 55 g 125 t

BASE COUNT

ORIGIN

Query Match 31.3%; Score 31.6; DB 9; Length 435;
 Best Local Similarity 62.8%; Pred. No. 3.9e+02;
 Matches 49; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 12 GTGCTCTGCTGACTCATTAGAGCTGGATGCTTTCTCTGTTGTAATCTTCTGTT 71
 DQ 1.1 GTGCTCTGCTGACTCATTAGAGCTGGATGCTTTCTCTGTTGTAATCTTCTGTT 71

DQ 285 GAGGCTAGAACCAATGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 226
 QY 72 CTTCAATAGATATGTAAG 89
 DQ 1.1 CTTCAATAGATATGTAAG 89

DQ 225 ATTATTTGATATATAGA 208
 QY 72 CTTCAATAGATATGTAAG 89
 DQ 1.1 CTTCAATAGATATGTAAG 89

RESULT 11

AQ651385

LOCUS

DEFINITION

ACCESSION

555 bp DNA linear GSS 22-JUN-1999

Sheared DNA-17L3.TR Sheared DNA Trypanosoma Brucei genomic clone

Sheared DNA-17L3, DNA sequence.

AQ651385

VERSION A0651385.1 GI:5144571
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei

REFERENCE 1 (bases 1 to 555)
AUTHORS El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Suh, E., Malek, J., Fujii, C., Gerrard, C., Leach, V., de Jong, P., Ullu, E., Melville, S., Donelson, J., Fraser, C., and Adams, M.

TITLE Determination of clone end sequences from Trypanosoma brucei GUTat 10.1 sheared DNA library

JOURNAL Unpublished (1999)

COMMENT Other GSSs: Sheared DNA-17L3-TF
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: neisayed@tigr.org

Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared DNA library constructed at TIGR. Clones will be available for distribution through ATCC. Sheared DNA end sequences search page: <http://www.tigr.org/cdb/mbd/tbdb/>.

Seq primer: M13-Reverse

Class: shotgun.

Location/Qualifiers

1. 555

/organism="Trypanosoma brucei"

/strain="TREU927/4 GUTat 10.1"

/db_xref="taxon:5691"

/clone="Sheared DNA-17L3"

/clone_lib="Sheared DNA"

/note="Vector: pUC18; Site 1: SmaI; Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (approx 2 kb). The v-i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borel, Oxford University Press, 1999)."

BASE COUNT 207 a 83 c 85 g 180 t

ORIGIN

Query Match 31.3%; Score 31.6; DB 17; Length 555;

Best Local Similarity 62.8%; Pred. No. 3.8e+02;

Matches 49; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 18 TCCTGACTCATTAGAGTGGAGCCCTTTCTCTGCTGCTGATTAATCTTTCTGTTTCAT 77

Db 465 TCCTGCTCAGCTGTCGCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTC 524

QY 78 TATATATGTAATGCTTT 95

Db 525 TAAATTTATGATGCTTT 542

RESULT 12

AL514767/c

LOCUS AL514767 LTI_NFL006.PL2 Homo sapiens cDNA clone CL08B0152G02 3

DEFINITION prime, mRNA sequence.

ACCESSION AL514767

VERSION AL514767.1 GI:12778260

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 578)

AUTHORS Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 578

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CL08B0152G02"

/clone_lib="LTI_NFL006.PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 150 a 119 c 135 g 138 t

ORIGIN

Query Match 31.3%; Score 31.6; DB 9; Length 578;

Best Local Similarity 58.1%; Pred. No. 3.8e+02;

Matches 43; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 17 CTCCTGACTCATTAGAGTGGAGCCCTTTCTCTGCTGCTGATTAATCTTTCTGTTTCAT 76

Db 150 CTCCTGACTCATTAGAGTGGAGCCCTTTCTCTGCTGCTGATTAATCTTTCTGTTTCAT 91

QY 77 TTAGATACTGTAAT 90

Db 50 TTTTCTTGTGAAT 77

RESULT 13

AQ8959744

LOCUS HS_5234.AL.F10.T7A RPI-11 Human Male BAC Library Homo sapiens

DEFINITION genomic clone Plate-9002 Col-19 Row-K, DNA sequence.

ACCESSION AQ8959744

VERSION AQ8959744.1 GI:63555934

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 911)

AUTHORS Mahairas, G.C., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.

TITLE Sequence-tagged connectors. A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98105, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPI-11. For BAC library availability, please contact Pletier de Jong

(pletier@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

or from Research Genetics (info@resgen.com). BAC end web server:

Hatches 51; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2 TATGTCCTGTGCTCTCCGACTCATATGAGCTGGATGCCCTTTCCCTGCTCTGATTAAT 61

Db 339 TTCTCTGATTGTCAGCATCAGAGTTTCACTCGTGGTGCATTTTCTCTCTGGCAAT 280

QY 62 CTTCCTGTTTCTTCATTAGATATGT 86

Db 279 CTTTGTGTACACCCTAATAATAT 255

Search completed: December 8, 2002, 07:50:31
for time : 790.132 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:10:23 ; Search time 26.618 Seconds
(without alignments)
1478.802 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600

Perfect score: 101
Sequence: 1 atgttgctgtgtgtctcc.....tatgtaaatgcttcaagt 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 350425 seqs, 194966369 residues

cal number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
C 1	32	31.7	288	10	US-09-815-242-4447
C 2	32	31.7	303	10	US-09-815-242-8393
C 3	32	31.7	303	10	US-09-815-242-8767
C 4	32	31.7	303	10	US-09-815-242-8943
C 5	30.6	30.3	3931	12	US-10-051-952-5
C 6	29.2	28.9	429	9	US-10-001-887-73
C 7	28.6	28.3	75899	10	US-09-854-883-243
C 8	28	27.7	5000	10	US-09-864-761-10713
C 9	28	27.7	5000	10	US-09-791-105-2
C 10	28	27.7	76798	10	US-09-880-107-3949
C 11	28	27.7	397658	10	US-09-813-320-3
C 12	27.8	27.5	4797	10	US-09-751-797-25
C 13	27.4	27.1	397658	10	US-09-813-320-3
C 14	27.2	26.9	21045	10	US-09-764-864-1695
C 15	27	26.7	147309	10	US-09-742-312-3
C 16	27	26.7	1503941	9	US-09-946-807-1
C 17	27	26.7	1503841	10	US-09-795-668-1
C 18	27	26.7	1503841	10	US-09-795-668-1
C 19	26.8	26.5	432	10	US-09-864-761-33271

20	26.8	26.5	555	10	US-09-864-761-16805
21	26.8	26.5	24533	9	US-09-764-868-1349
C 22	26.6	26.3	387	10	US-09-864-761-10466
C 23	26.6	26.3	2000	9	US-09-938-842A-5172
C 24	26.6	26.3	2529	10	US-09-764-847-1859
C 25	26.6	26.3	17569	9	US-09-736-457-1804
C 26	26.6	26.3	17569	9	US-09-902-941-1804
C 27	26.6	26.3	115592	10	US-09-818-512-3
C 28	26.4	26.1	1805	10	US-09-728-422-7
C 29	26.4	26.1	2724	10	US-09-511-781-31
C 30	26.4	26.1	3164	10	US-09-764-869-1451
C 31	26.4	26.1	3571	9	US-10-135-583-42
C 32	26.4	26.1	3571	10	US-09-823-033-3
C 33	26.4	26.1	4500	10	US-09-729-653-1
C 34	26.4	26.1	14171	10	US-09-764-877-3517
C 35	26.4	26.1	15958	10	US-09-764-877-3518
C 36	26.2	25.9	2309	10	US-09-925-300-680
C 37	26.2	25.9	2457	10	US-09-739-254-40
C 38	26.2	25.9	2457	10	US-09-904-615-40
C 39	26.2	25.9	38374	10	US-09-880-107-3463
C 40	26.2	25.9	249467	9	US-10-026-188-3
C 41	26	25.7	347	10	US-09-783-590-7481
C 42	26	25.7	450	10	US-09-834-975-622
C 43	26	25.7	450	10	US-09-834-975-762
C 44	26	25.7	656	10	US-09-878-574-4414
C 45	26	25.7	1381	10	US-09-925-301-276

ALIGNMENTS

RESULT 1
US-09-815-242-4447/C
; Sequence 4447, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Irawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4447
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4447

Query Match 31.7% Score 32; DB 10; Length 288;
Best Local Similarity 66.8%; Pred. No. 2.2;

Db 1656 TGAT 1659

RESULT 15

US-09-742-312-3/c
; Sequence 3, Application US/09742312
; Patent No. US20020045166A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000838
; CURRENT APPLICATION NUMBER: US/09/742,312
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 147309
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(147309)
; OTHER INFORMATION: n = A,T,C or G
US-09-742-312-3

Query Match 26.7%; Score 27; DB 10; Length 147309;
Best Local Similarity 66.1%; Pred. No. 1.8e+02;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 23 ACTCATTAGAGCTGGATGCTTTTCCTGCTGTGATAATCTTTCTGTTCTTCATTAGA 81
Db 49881 AATCATTAGTGTAGTGTCCAAATAGCCCGGAGTATGCTTTCTGGTCTTGTGA 49823

Search completed: December 8, 2002, 09:29:51
Job time : 274.632 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 8, 2002, 07:52:47 : Search time 5277.36 Seconds
(without alignments)
11586.276 Million cell updates/sec

Title: X51602_COPY_1900_4000
Perfect score: 2101
Sequence: 1 TATATCAGAGATGTCACAA.....CCACAGCAGCAGTCGTCG 2101

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
4109280
tal number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenBank:
1: gb-be:	
2: gb-hgt:	
3: gb-in:	
4: gb-om:	
5: gb-ov:	
6: gb-pat:	
7: gb-ph:	
8: gb-pl:	
9: gb-pr:	
10: gb-ro:	
11: gb-sts:	
12: gb-sy:	
13: gb-un:	
14: gb-vi:	
15: em-ba:	
16: em-fun:	
17: em-hum:	
18: em-in:	
19: em-mu:	
20: em-om:	
21: em-or:	
22: em-ov:	
23: em-pat:	
24: em-ph:	
25: em-fl:	
26: em-fo:	
27: em-sts:	
28: em-un:	
29: em-vi:	
30: em-hgt_hum:	
31: em_hgt_in:	
32: em_hgt_other:	
33: em_hgt_mus:	
34: em_hgt_pln:	
35: em_hgt_rod:	
36: em_hgt_mam:	
37: em_hgt_vrt:	
38: em_sy:	
39: em_hgtgo_hum:	
40: em_hgtgo_mus:	
41: em_hgtgo_other:	

Result No.	Score	Match	Length	DB ID	Description
1	2101	100.0	7680	9	HSFLT Human flt m
2	2097.8	99.8	4017	6	AX481481 Sequence
3	2097.8	99.8	4017	9	AF063657 Homo sapi
4	1599.2	76.1	4734	10	D28498 Rat mRNA fo
5	1574.5	74.9	6055	10	L07297 Mus musculu
6	1574.5	74.9	6055	10	D88689 Mus musculu
7	1534.2	73.0	5482	10	M88689 Mus musculu
8	1168.8	55.6	4272	5	AB065372 Gallus ga
9	956.4	45.5	22065	2	AB065372 Gallus ga
10	953.2	45.4	197279	9	AC104439 Homo sapi
11	713.8	34.0	4044	5	AX481480 Sequence
12	713.8	34.0	4071	6	AR142803 Sequence
13	713.8	34.0	4071	6	AR201382 Sequence
14	713.8	34.0	4071	6	AX481483 Sequence
15	713.8	34.0	4071	9	AF063658 Homo sapi
16	713.8	34.0	4225	6	AX133539 Sequence
17	713.8	34.0	5830	9	AF035121 Homo sapi
18	710.6	33.8	4230	9	HSFLT Homo sapien
19	710.6	33.8	4230	9	HUNKDRZ Sequence
20	710.6	33.8	4236	5	AR030758 Sequence
21	710.6	33.8	5464	10	BC020530 Mus muscu
22	710.6	33.8	5470	6	AR068047 Sequence
23	710.6	33.8	5470	6	AR125598 Sequence
24	710.6	33.8	5470	10	MMFLKIM Sequence
25	710.6	33.8	5470	10	MMFLK1 Mouse Flk-1
26	708.4	33.7	5391	10	MMFLK1 Sequence
27	708.4	33.7	5406	6	AR005213 Sequence
28	708.4	33.7	5406	6	AR071705 Sequence
29	708.4	33.7	5406	6	125171 Sequence 5
30	708.4	33.7	5406	6	140602 Sequence 5
31	706.8	33.6	4790	10	S53103 NKX/FLK-1=
32	693.2	33.0	5390	5	CCQUEK1 C.coturnix
33	690	32.8	2352	6	AR030849 Sequence 17
34	690	32.8	2352	6	182808 Sequence 17
35	665.8	31.7	5037	5	AF180354 Danio rer
36	661	31.5	3906	5	AY056466 Danio rer
37	640.6	30.5	4617	5	CCQUEK2 C.coturnix
38	624	29.7	2523	6	AR163540 Sequence
39	624	29.7	2523	6	E13256 Human mRNA
40	624	29.7	2523	6	E14000 Human mRNA
41	609.6	29.0	3009	6	AX060540 Sequence
42	561.8	26.7	5284	10	M88689 Mus musculu
43	541.6	25.8	3897	6	AX481482 Sequence
44	541.6	25.8	4195	6	AR016568 Sequence
45	541.6	25.8	4195	6	AR106405 Sequence

ALIGNMENTS

Result 1	HSFLT	Human flt mRNA for receptor-related tyrosine kinase.	7680 bp	mRNA	linear	PRI 15-NOV-1993
LOCUS	HSFLT	Human flt mRNA for receptor-related tyrosine kinase.	7680 bp	mRNA	linear	PRI 15-NOV-1993
DEFINITION	HSFLT	Human flt mRNA for receptor-related tyrosine kinase.	7680 bp	mRNA	linear	PRI 15-NOV-1993
ACCESSION	X51602	GI:31431				
VERSION	X51602.1	GI:31431				
KEYWORDS	flt gene; fms-related tyrosine kinase gene; tyrosine kinase.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 7680)					
AUTHORS	Shibuya M.					
TITLE	Direct Submission					
JOURNAL	Submitted (02-JAN-1989) Shibuya M., Institute of Medical Science,					

Pred. No. is the number of results predicted by chance to have a

QY 1261 TTTCAGAGATATAAGTCTGATGTTGAGAGAGAGAGATCTGACGGTTTCTAC 1320
Db 3160 TTTCAGAGATATAAGTCTGATGTTGAGAGAGAGAGATCTGACGGTTTCTAC 3219
QY 1321 AAGAGGCCATCACTATGAAGATCTGATTTTACAGTTTTCAGTGGTCCAGAGGCAIG 1380
Db 3220 AAGAGGCCATCACTATGAAGATCTGATTTTACAGTTTTCAGTGGTCCAGAGGCAIG 3279
QY 1381 GAGTCTGCTCTTCCAGAAAGTGCATTCAGTGGTCCAGAGGCAAGATTTCTTTA 1440
Db 3280 GAGTCTGCTCTTCCAGAAAGTGCATTCAGTGGTCCAGAGGCAAGATTTCTTTA 3339
QY 1441 TCTGAGACAGCTGGTGGAGATTTGATTTTGGCTTCCGCGGATATTTATAGAAC 1500
Db 3340 TCTGAGACAGCTGGTGGAGATTTGATTTTGGCTTCCGCGGATATTTATAGAAC 3399
QY 1501 CCCGATTAATGAGAAAGGAGATCTGAGTCTCTGAAATGATGGTCCCGAATCI 1560
Db 3400 CCCGATTAATGAGAAAGGAGATCTGAGTCTCTGAAATGATGGTCCCGAATCI 3459
QY 1561 ATCTTTCAGAAATCTACGACCAAGAGAGAGAGTGGTCTTACGAGTATTTCTGG 1620
Db 3460 ATCTTTCAGAAATCTACGACCAAGAGAGAGAGTGGTCTTACGAGTATTTCTGG 3519
QY 1621 GAATCTTCTCTTAGTGGTCTCCATACCCAGAGATCAAAATGATGATGATTTGG 1680
Db 3520 GAATCTTCTCTTAGTGGTCTCCATACCCAGAGATCAAAATGATGATGATTTGG 3579
QY 1681 AGTCCCTGAGGAGGAGATGAGTGGTCTCTGAGTCTCTGAGTCTCTGAGTCTCT 1740
Db 3580 AGTCCCTGAGGAGGAGATGAGTGGTCTCTGAGTCTCTGAGTCTCTGAGTCTCT 3639
QY 1741 CAGATCATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1800
Db 3640 CAGATCATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 3699
QY 1801 GTGGAATAATAGTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1860
Db 3700 GTGGAATAATAGTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3759
QY 1861 CCATCATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1920
Db 3760 CCATCATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 3819
QY 1921 GAGGATCTTTCAGGAAAGTATTTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 1980
Db 3820 GAGGATCTTTCAGGAAAGTATTTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 3879
QY 1981 GTGAGATATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2040
Db 3880 GTGAGATATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3939
QY 2041 CTCTTACGATGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 2100
Db 3940 CTCTTACGATGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 3999
QY 2101 G 2101
Db 4000 G 4000

RESULT 2
AX481481 4017 bp DNA linear PAT 16-AUG-2002
LOCUS
DEFINITION Sequence 95 from Patent WO02055693.
ACCESSION AX481481
VERSION AX481481.1 GI:2316395
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.
TITLE Method for inhibiting the expression of a target gene
JOURNAL Patent: WO 02055693-A 95 18-JUL-2002;
Ribopharma AG (DE)
FEATURES Location/Qualifiers
Source 1. 4017
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1236 a 915 c 903 g 963 t
ORIGIN
Query Match 99.8%; Score 2097.8; DB 6; Length 4017;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2099; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TATATCAGATGTCCTAATGGTTTTCATGTTAACTTGGAAAAATCCGACGGAGGA 60
Db 1551 TATATCAGATGTCCTAATGGTTTTCATGTTAACTTGGAAAAATCCGACGGAGGA 1710
QY 51 GAGGACCTGAACTGCTTGGACAGTTTAAAGTTCTTATACAGAGAGCTTTACTTGA 120
Db 1711 GAGGACCTGAACTGCTTGGACAGTTTAAAGTTCTTATACAGAGAGCTTTACTTGA 1770
QY 121 TTACTGGGACAGTTTAAACAGACCAATGCTACAGTATAGCAAGCAAAATGGCC 180
Db 1771 TTACTGGGACAGTTTAAACAGACCAATGCTACAGTATAGCAAGCAAAATGGCC 1830
QY 181 ATCACTAAGGAGGACCTCCATCACTCTTAACTTACCATCATGAATGTTTCCCTCCA 240
Db 1831 ATCACTAAGGAGGACCTCCATCACTCTTAACTTACCATCATGAATGTTTCCCTCCA 1890
QY 241 TCAGGACCTTATGCTTGCAGAGCCAGGAATGTATACAGGGGAGAAATCTCTCCAGA 300
Db 1891 TCAGGACCTTATGCTTGCAGAGCCAGGAATGTATACAGGGGAGAAATCTCTCCAGA 1950
QY 301 AAGAAATTAACATCAGATCAGAGACCAACATACCTCTGCGAAACCTCAGTGATC 360
Db 1951 AAGAAATTAACATCAGATCAGAGACCAACATACCTCTGCGAAACCTCAGTGATC 2010
QY 361 ACAGTGGCCATCAGAGTCTCCACCTTTAGACTGTCTAATAGTGTGCTCCCGAGCT 420
Db 2011 ACAGTGGCCATCAGAGTCTCCACCTTTAGACTGTCTAATAGTGTGCTCCCGAGCT 2070
QY 421 CAGATCATCTGTTTAAACACCAACCAATACACAGAGCCCTGGAATTTATTTAG 480
Db 2071 CAGATCATCTGTTTAAACACCAACCAATACACAGAGCCCTGGAATTTATTTAG 2130
QY 481 CCAGGAGGACGACGCTGTTTATGAAAGTTCACAGAGGATGAAAGTGTCTATC 540
Db 2131 CCAGGAGGACGACGCTGTTTATGAAAGTTCACAGAGGATGAAAGTGTCTATC 2190
QY 541 TCGAAGCCACCAACCAAGGCTCTGCGAAGTTCAGCATCTCAGTCTGTTTCAAG 500
Db 2191 TCGAAGCCACCAACCAAGGCTCTGCGAAGTTCAGCATCTCAGTCTGTTTCAAG 2250
QY 501 ACCCTGGACAGGCTTAACTGAGGCTGATCACTCTAATGAGGCTGCTGCTGGGAC 560
Db 2251 ACCCTGGACAGGCTTAACTGAGGCTGATCACTCTAATGAGGCTGCTGCTGGGAC 2310
QY 561 CTCTTCTGGTCTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 720
Db 2311 CTCTTCTGGTCTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 2370
QY 721 AAGACTGACCTATCAATTAATGAGGCTGAGTCTGCTTGGATGAGCAGTGT 780
Db 2371 AAGACTGACCTATCAATTAATGAGGCTGAGTCTGCTTGGATGAGCAGTGT 2430
QY 781 GAGGCTCTCTTATGATCCAGAGTGGGAGTTTCCCGGGAGAGACTTAACTGGG 840
Db 2431 GAGGCTCTCTTATGATCCAGAGTGGGAGTTTCCCGGGAGAGACTTAACTGGG 2490
QY 841 AAATCATTGGAAGAGGCTTTTGGAAAGTGTTCAGCATCAGCAATTTGGCAAT 900

FAELVEKLGLDLOANVQDQKDIPIINAILTNGSGFTVSTPFAFDEFFKESISAPKEN
SGSDVRVNAFKMSLERIKIFELLFNATSMEDFYQSDSSTLLASPMKRFETWD
SKFKSLKIDLKVIKSKESGLSDVSRPSCFSSCHVSEGRKRFITYDHAELERKTIAC
CSPPDYNSWVLYSTPPI*
BASE COUNT 1236 a 915 c 903 g 963 t

Query Match		99.8%	Score 2097.8	DB 9	Length 4017
Best local Similarity		99.9%	Pred. No. 0		
Matches 2099		Conservative 0	Mismatches 2	Indels 0	Gaps 0
Oy	1	TATATCAGAGATGCGCAAGTGGTTCATGTATACCTTGGAAATAATGCCGACGAAGA	60		
Db	1651	TATATCAGAGATGCGCAAGTGGTTCATGTATACCTTGGAAATAATGCCGACGAAGA	1710		
Oy	51	GAGGACCTGAACATGCTCTTCCACAGTATACAAAGTCTTATACAGACAGCTTACCTGGATT	120		
Db	1711	GAGGACCTGAACATGCTCTTCCACAGTATACAAAGTCTTATACAGACAGCTTACCTGGATT	1770		
Oy	121	TTACTGCGGACAGTATATACAGCAATATGCTATACCTATACAGTATATACAGCAATATGCTT	180		
Db	1771	TTACTGCGGACAGTATATACAGCAATATGCTATACCTATACAGTATATACAGCAATATGCTT	1830		
Oy	181	ATCAGTAAAGGACAGCTCCATCAGCTTATATATACATCATGAAGTTCCTTGCAGAT	240		
Db	1831	ATCAGTAAAGGACAGCTCCATCAGCTTATATATACATCATGAAGTTCCTTGCAGAT	1890		
Oy	241	TCAGGACCTATGCTGCGAGCGCAGGAATGTATACAGGGGAGGAATATCTTCAGAG	300		
Db	1891	TCAGGACCTATGCTGCGAGCGCAGGAATGTATACAGGGGAGGAATATCTTCAGAG	1950		
Oy	301	AAAGAAATTTACATCAGAGATCAGGAGCAGCATACCTCTCTCGAAGACCTCAGTATCAC	360		
Db	1951	AAAGAAATTTACATCAGAGATCAGGAGCAGCATACCTCTCTCGAAGACCTCAGTATCAC	2010		
Oy	361	ACAGTGGCCATCAGCAGTCCACACCTTTAGACTGTCAAGTATAGTGTTCCTCCGAGCT	420		
Db	2011	ACAGTGGCCATCAGCAGTCCACACCTTTAGACTGTCAAGTATAGTGTTCCTCCGAGCT	2070		
Oy	421	CAGATCAGTGGTATTAACACACCAACAAATACACAGAGAGCTGCAATATTTTNGA	480		
Db	2071	CAGATCAGTGGTATTAACACACCAACAAATACACAGAGAGCTGCAATATTTTNGA	2130		
Oy	481	CCAGGAGCAGCAGCTGTTTATGAAGATCAGAGAGAGGATGAAGTGTCTATCAC	540		
Db	2131	CCAGGAGCAGCAGCTGTTTATGAAGATCAGAGAGAGGATGAAGTGTCTATCAC	2190		
Oy	541	TGCAAGGACCAACACAGAGGCTGTGTGGAAGTTCAGCATACCTCAGTGTCAAGGA	600		
Db	2191	TGCAAGGACCAACACAGAGGCTGTGTGGAAGTTCAGCATACCTCAGTGTCAAGGA	2250		
Oy	601	ACCTGGGACAGTATATCTGGAGCTGATCAGCTATACATGCACTGTGTGGCTGGACT	660		
Db	2251	ACCTGGGACAGTATATCTGGAGCTGATCAGCTATACATGCACTGTGTGGCTGGACT	2310		
Oy	661	CTCTTCTGGCTCTTATTAACCTCTCTTATCCGAAATATGAAGGCTCTCTCTGAATA	720		
Db	2311	CTCTTCTGGCTCTTATTAACCTCTCTTATCCGAAATATGAAGGCTCTCTCTGAATA	2370		
Oy	721	AGAGCTAGCTATCTATATATATGAGCCAGAGTGAAGTCTTCTTGGATGAGCAGTGT	780		
Db	2371	AGAGCTAGCTATCTATATATATGAGCCAGAGTGAAGTCTTCTTGGATGAGCAGTGT	2430		
Oy	781	GAGCGCTCCCTTATGATGCCAGCAGTGGAGTGTGCGGCGAGACATTAACCTGGCC	840		
Db	2431	GAGCGCTCCCTTATGATGCCAGCAGTGGAGTGTGCGGCGAGACATTAACCTGGCC	2490		
Oy	841	AAATACCTTGGAGAGGGCTTTTGGAAAGTGGTTCAGGATCAGCATTTGGCATTTAAG	900		
Db	2491	AAATACCTTGGAGAGGGCTTTTGGAAAGTGGTTCAGGATCAGCATTTGGCATTTAAG	2550		
Oy	901	AAATACCTTGGAGAGGGCTTTTGGAAAGTGGTTCAGGATCAGCATTTGGCATTTAAG	960		

Db	2551	AAATCAGCTATGCTCCGAGACTGTGGCTGTGAAAAATGCTGAAGAGGGGCGCCAGCCGACG	2610		
Oy	961	GAGTACAAAGCTCTCATGACTAGCTATAAATCTTTGACCCACATTTGCCACCATCTGAAC	1020		
Db	2611	GAGTACAAAGCTCTCATGACTAGCTATAAATCTTTGACCCACATTTGCCACCATCTGAAC	2670		
Oy	1021	GTGGTATACCTGCTGGAGCCTGCACCAAGCAAGAGGGCTCTGTATGTGATTTGAA	1080		
Db	2671	GTGGTATACCTGCTGGAGCCTGCACCAAGCAAGAGGGCTCTGTATGTGATTTGAA	2730		
Oy	1081	TACTGCAATATGAAATCTCTCCAACTACCTCAAGAGCAAGCTGACTTATTTTCTC	1140		
Db	2731	TACTGCAATATGAAATCTCTCCAACTACCTCAAGAGCAAGCTGACTTATTTTCTC	2790		
Oy	1141	AAACAAGGATGACGACCTACACATCGAGCCTTAAGAAGAAAAAATGGAGCCCTGGAA	1200		
Db	2791	AAACAAGGATGACGACCTACACATCGAGCCTTAAGAAGAAAAAATGGAGCCCTGGAA	2850		
Oy	1201	CAAGSCAAGAAACCAAGACTAGTAGTNGCTTACCACGACGCAAGACTTTGGAGCTCCGGC	1260		
Db	2851	CAAGSCAAGAAACCAAGACTAGTAGTNGCTTACCACGACGCAAGACTTTGGAGCTCCGGC	2910		
Oy	1261	TTTCAGGAAGATATAAGTCTGAGTGTATGTTGAGGAAGAGGAGGATCTGACGGTTTCTAC	1320		
Db	2911	TTTCAGGAAGATATAAGTCTGAGTGTATGTTGAGGAAGAGGAGGATCTGACGGTTTCTAC	2970		
Oy	1321	AGAGGCCCATCTACTATGGAAGATCTGATTTCTTACAGTTTTCAGTGGCCAGAGGCAATG	1380		
Db	2971	AGAGGCCCATCTACTATGGAAGATCTGATTTCTTACAGTTTTCAGTGGCCAGAGGCAATG	3030		
Oy	1381	GAGTCTCTGCTTCCAGAAAGTGCATTCATCGGACCTTCCAGGAGCTTGGAGGAGAAACATCTTTT	1440		
Db	3031	GAGTCTCTGCTTCCAGAAAGTGCATTCATCGGAGCTTGGAGGAGAAACATCTTTT	3090		
Oy	1441	TCTGAGAACACGCTGGTGAAGATTTTGTGATTTTGGCTTTGCCCTTGGCCGGATATTTAAGAAC	1500		
Db	3091	TCTGAGAACACGCTGGTGAAGATTTTGTGATTTTGGCTTTGCCCTTGGCCGGATATTTAAGAAC	3150		
Oy	1501	CCCGATTTGTGAGAAAGAGAGATCTCGACATTCCTCTGAAATGGATGCTCCGAACTCT	1560		
Db	3151	CCCGATTTGTGAGAAAGAGAGATCTCGACATTCCTCTGAAATGGATGCTCCGAACTCT	3210		
Oy	1561	ATCTTTGACAAATCTACAGCACCAGAGCGAGTGTGGTCTTACGAGATTTGCTGTGG	1620		
Db	3211	ATCTTTGACAAATCTACAGCACCAGAGCGAGTGTGGTCTTACGAGATTTGCTGTGG	3270		
Oy	1621	GAAATCTCTCTTAGTGGGTCTCCATACCCAGGATACAAATGGAATGAGGACTTTTGC	1680		
Db	3271	GAAATCTCTCTTAGTGGGTCTCCATACCCAGGATACAAATGGAATGAGGACTTTTGC	3330		
Oy	1681	AGTCCCTGAGGAGGAGCATGAGGATGAGAGCTCTGAGTACTCTACTCTGAAATCTAT	1740		
Db	3331	AGTCCCTGAGGAGGAGCATGAGGATGAGAGCTCTGAGTACTCTACTCTGAAATCTAT	3390		
Oy	1741	CAGATCACTGGACTGTGSCACAGAGACCCCAAGAAAGGCCCAAGATTTGAGAACTT	1800		
Db	3391	CAGATCACTGGACTGTGSCACAGAGACCCCAAGAAAGGCCCAAGATTTGAGAACTT	3450		
Oy	1801	GTGGAAGAACTAGGTGATTTGCTTCAAGCAATATGACCAAGAGTGTGAAGACTACATC	1860		
Db	3451	GTGGAAGAACTAGGTGATTTGCTTCAAGCAATATGACCAAGAGTGTGAAGACTACATC	3510		
Oy	1861	CCAAATCAATGCCATACCTGACAGGAATATGAGTGTATACATCACTCACTCTGCTTCTCT	1920		
Db	3511	CCAAATCAATGCCATACCTGACAGGAATATGAGTGTATACATCACTCACTCTGCTTCTCT	3570		
Oy	1921	GAGGACTCTTCAAGAAAGTATTTAGCTTCCGAAGTTTAAATTCAGGAAGCTCTGATGAT	1980		
Db	3571	GAGGACTCTTCAAGAAAGTATTTAGCTTCCGAAGTTTAAATTCAGGAAGCTCTGATGAT	3630		
Oy	1981	GTGAGATATGTAATGCTTCAAGTTCATGAGCTTGGAAAGAAATCAAAACCTTTGAGAA	2040		
Db	3631	GTGAGATATGTAATGCTTCAAGTTCATGAGCTTGGAAAGAAATCAAAACCTTTGAGAA	3690		

polyA_site

DYNWVLYSSPPA*

5055

/sere="FLI"

BASE COUNT 1633 a 1549 c 1492 g 1381 t

ORIGIN

Query Match 74.9%; Score 1574.6; DB 10; Length 6055;
Best Local Similarity. 84.7%; Pred. No. 0;
Matches 1779; Conservative 0; Mismatches 319; Indels 3; Gaps 1;

QY 1 TATATACAGATGTCGCAATGCGGTTTCATGTTAACTTGGAAAAATGCCGAGCGAAGA 60
DB 1680 TATGTCACAGATGTCGCAATGCGGTTTCATGTTAACTTGGAAAAATGCCGAGCGAAGA 739
QY 61 GAGGACCTCAAGTGTCTGACAGTATACAGATCTTATACAGAGCGTTTACTTGGAT 120
DB 1740 GAGGACCTCAAGTGTCTGACAGTATACAGATCTTATACAGAGCGTTTACTTGGAT 1799
QY 121 TTAATCGGACAGTATATACAGACAAATGCCACTACAGTATATACAGAGCGTTTACTTGGAT 180
DB 1800 CTGCTACGACAGTATATACAGACAAATGCCACTACAGTATATACAGAGCGTTTACTTGGAT 1859
QY 181 ATCACTAGGACACTGCATCTTATATCTTATCACTATGATGTTTCCCTGCAAGAT 240
DB 1860 ACCACTCAAGATATCACTCTCACTTCACTTGTATCAAGAACGTTCTCTACAGAC 1919
QY 241 TCAGGACCTATGCTGACAGCGAGGATGTATACAGAGCGGAGAGATCTCTCAGAG 300
DB 1920 TCGGACCTATGCTGACAGCGAGGATGTATACAGAGCGGAGAGATCTCTCAGAG 1979
QY 301 AAGAAATATCAATACAGACATGACAGGACCACTATCTCTGGAACCTCAGTATAC 360
DB 1980 ACAGAGTCTCGTTAGAGATTCGGAAGCGCCACAGCTGCTCANACCTCAGTATAC 2039
QY 361 ACAGTGGCCATCAGAGTCCACCTTTAGAGTGTCACTGTTATGTTGTTCCCGAGCT 420
DB 2040 GAGTCTCACTAGGCTCTAGAGCTTTAGAGTGTCACTGTTATGTTGTTCCCGAGCT 2099
QY 421 CAGATCACTTGGTTTAAACACACCAATATACAGAGGCTTGAATATATTTAGGA 480
DB 2100 CAGATCACTTGGTTTAAACACACCAATATACAGAGGCTTGAATATATTTAGGA 2159
QY 481 CCAGAGACAGCGCTGTTTATGAAAGAGTACAGAGGATGAGGTTGTTCTATAC 540
DB 2160 CCAGAGACAGCGCTGTTTATGAAAGAGTACAGAGGATGAGGTTGTTCTATAC 2219
QY 541 TGAAGACAGCGCTGTTTATGAAAGAGTACAGAGGATGAGGTTGTTCTATAC 600
DB 2220 TGAAGACAGCGCTGTTTATGAAAGAGTACAGAGGATGAGGTTGTTCTATAC 2279
QY 601 ACCTCGGACAGTCTAACTGAGCTGATCACTTAACTGACCTGTGCTGCTGCGACT 660
DB 2280 ACCTCGGACAGTCTAACTGAGCTGATCACTTAACTGACCTGTGCTGCTGCGACT 2339
QY 661 CTCCTGCTGCTTATTAACCTCTTATCGGAAATGAAAGGTTCTTCTGAAATA 720
DB 2340 CTCCTGCTGCTTATTAACCTCTTATCGGAAATGAAAGGTTCTTCTGAAATA 2399
QY 721 AAGCTGACTACCTATCAATATATGAGGACCGATGAGTCTTCTGATGAGCTG 780
DB 2400 AAGCTGACTACCTATCAATATATGAGGACCGATGAGTCTTCTGATGAGCTG 2459
QY 781 GAGGCTGCTTATGATGCGAGAGTGGAGTTGCGGCGAGAGACTTAACTGCGC 840
DB 2460 GAGGCTGCTTATGATGCGAGAGTGGAGTTGCGGCGAGAGACTTAACTGCGC 2519
QY 841 AAATCACTTGGAGAGGCGCTTTTGGAAAGTGGTTTCAGCGGAGAGACTGNACTAGC 900
DB 2520 AAATCACTTGGAGAGGCGCTTTTGGAAAGTGGTTTCAGCGGAGAGACTGNACTAGC 2579
QY 901 AAATCACTTGGAGAGGCGCTTTTGGAAAGTGGTTTCAGCGGAGAGACTGNACTAGC 960
DB 2580 AAATCACTTGGAGAGGCGCTTTTGGAAAGTGGTTTCAGCGGAGAGACTGNACTAGC 2639

QY 961 GAGTCAARAGCTCTGATGACTGAGCTAGCTTAAATACTTGACCCACATTTGGCCACCATCTGAAC 1020
DB 2640 GAGTCAARAGCTCTGATGACTGAGCTAGCTTAAATACTTGACCCACATTTGGCCACCATCTGAAC 2699
QY 1021 GTGGTTAACTCTGCTGGAGCTGACCAAGAGGAGGCTCTGATGGTATGTTGAA 1080
DB 2700 GTGGTTAACTCTGCTGGAGCTGACCAAGAGGAGGCTCTGATGGTATGTTGAA 2759
QY 1081 TACTGCAATATGGAATCTCTCAACTTACCTCAAGAGCAAACTGACTTATTTCTC 1140
DB 2760 TACTGCAATATGGAATCTCTCAACTTACCTCAAGAGCAAACTGACTTATTTCTC 2819
QY 1141 AACAGAGTGGAGCTTACATGAGGCTTAAAGAGAAATGAGGAGGCTGAGG 1200
DB 2820 AACAGAGTGGAGCTTACATGAGGCTTAAAGAGAAATGAGGAGGCTGAGG 2879
QY 1201 CAAGGCAAGAAACCAAGACTAGATAGCTTACAGGACGCGAAAGCTTTGGAGCTCCGGC 1260
DB 2880 CAAGGCAAGAAACCAAGACTAGATAGCTTACAGGACGCGAAAGCTTTGGAGCTCCGGC 2936
QY 1261 TTTGAGGAGATTAAGTCTGAGTGAITGTAGGAGAGGAGGATTTGACGGTTTCTAC 1320
DB 2937 TTTGAGGAGATTAAGTCTGAGTGAITGTAGGAGAGGAGGATTTGACGGTTTCTAC 2996
QY 1321 AAGGACCCCTACATATGGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCATG 1380
DB 2997 AAGGACCCCTACATATGGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCATG 3056
QY 1381 GAGTCTCTCTTCCAGAAAGTGCATTCATGAGGACCTGGCAGAGAAACATTTCTTTA 1440
DB 3057 GAGTCTCTCTTCCAGAAAGTGCATTCATGAGGACCTGGCAGAGAAACATTTCTTTA 3116
QY 1441 TCTGAGAAACAGTGTGAGATTTGAGTTTGGCTTTCGCCGGGATATTTAAGAAC 1500
DB 3117 TCTGAGAAACAGTGTGAGATTTGAGTTTGGCTTTCGCCGGGATATTTAAGAAC 3176
QY 1501 CCGATATGAGGAAAGGAGATCTGACTTCCCTTAAATGAGTGGCTTCTGATTC 1560
DB 3177 CCGATATGAGGAAAGGAGATCTGACTTCCCTTAAATGAGTGGCTTCTGATTC 3236
QY 1561 ATCTTTGACAAATCTACAGCACCAGAGGAGGCTGCTTACGGAGTATTTCTGTTG 1620
DB 3237 ATCTTTGACAAATCTACAGCACCAGAGGAGGCTGCTTACGGAGTATTTCTGTTG 3296
QY 1621 GAATCTTCTCTTGGTCTCCATACCCAGAGTACAAATGATGAGGACTTTTC 1680
DB 3297 GAGATCTTCTCTTGGTCTCCATACCCAGAGTACAAATGATGAGGACTTTTC 3356
QY 1681 AGTCCCTGAGGAGGAGGATGAGGATGAGGCTCTGAGTACTTCTCTGAAATCTAT 1740
DB 3357 AGTCCCTGAGGAGGAGGATGAGGATGAGGCTCTGAGTACTTCTCTGAAATCTAT 3416
QY 1741 CAGATCACTGCTGAGTCTGCGACAGAGACCCAAAGAGGAGGAGGATTTGCAAGACT 1800
DB 3417 CAAATCACTGAGTCTGCGACAGAGACCCAAAGAGGAGGAGGATTTGCAAGACT 3476
QY 1801 GTGAAATCTGAGTCTGCGACAGAGGATGAGGATGAGGCTCTGAGTACTTCTGAA 1860
DB 3477 GTGAAATCTGAGTCTGCGACAGAGGATGAGGATGAGGCTCTGAGTACTTCTGAA 3536
QY 1861 CCAATCACTGAGTCTGCGACAGAGGATGAGGATGAGGCTCTGAGTACTTCTGAA 1920
DB 3537 CCAATCACTGAGTCTGCGACAGAGGATGAGGATGAGGCTCTGAGTACTTCTGAA 3596
QY 1921 GAGGACTTCTTCAAGGAAAGTATTTCACTTCCGAGGTTTAAATTCAGGAGGCTCTGAT 1980
DB 3597 GAGGACTTCTTCAAGGAAAGTATTTCACTTCCGAGGTTTAAATTCAGGAGGCTCTGAT 3656
QY 1981 GTCAGATATGTAATGCTTCAAGTTCATGAGGCTTGAAGAAATCAAACTTTGAGAA 2040
DB 3657 GTCAGATATGTAATGCTTCAAGTTCATGAGGCTTGAAGAAATCAAACTTTGAGAA 3716

```

QY 2041 CTTTACCGAATGCCACCTCCAGTTCAGTACACTACAGGCGAGCAGCAGCACTCTGTC 2100
||||| ||||||| ||||||| ||||||| ||| ||| ||||| ||||| |||
Db 3717 CTTTACCGAATGCCACCTCCAGTTCAGTACACTACAGTACAGCACTCTGTC 3776
QY 2101 G 2101
|
Db 3777 G 3777

RESULT 6
LOCUS D88689
DEFINITION Mus musculus mRNA for flt-1, complete cds.
ACCESSION D88689
VERSION D88689.1 GI:2809068
KEYWORDS
SOURCE Mus musculus (strain:C57/B) lung cdna to mRNA, clone_l1b,mouse lung
cDNA clone:MD-9-11.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Finnerty,H., Kelleher,K., Morris,G.E., Bean,K., Merberg,D.M.,
Kriz,R., Morris,J.C., Sookdeo,H., Turner,K.J. and Wood,C.R.
TITLE Molecular cloning of murine FLI and FLI4
MOLINE Oncogene 8 (8), 2293-2298 (1993)
PUBMED 9330572
REFERENCE 8393164
AUTHORS Kondo,K., Hiratuka,S., Subbalakshmi,E., Matsushima,H. and
Shibuya,M.
TITLE Genomic organization of the flt-1 gene encoding for Vascular
Endothelial growth factor (VEGF) receptor-1 suggests an intimate
evolutionary relationship between the 7-Ig and the 5-Ig tyrosine
kinase receptor
JOURNAL Gene (1998) In press
REFERENCE 3 (bases 1 to 6275)
AUTHORS Kondo,K.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-1996) Kunio Kondo, The Institute of Medical
Science, The University of Tokyo, Department of Clinical Oncology:
4-6-1 Shirokanedai, Minato-Ku, Tokyo 108, Japan
(Tel:-81-3-5449-5632, Fax:+81-3-5449-5428)
FEATURES
source
1..6275
/organism="Mus musculus"
/strain="C57/B"
/db_xref="taxon:10090"
/chromosome="5"
/map="5G1-G3"
/clone="MD-9-11"
/tissue_type="lung"
/clone_lib="mouse lung cdna"
1..6275
1..315
/number=1
/evidence=experimental
1..251
252..4253
/function="receptor for VEGF and PlGF"
/standard_name="fms-like tyrosine kinase 1"
/codon_start=1
/evidence=experimental
/product="flt-1"
/protein_id="BAA24498.1"
/db_xref="GI:2809068"
/translation="MYSQWDIAVLFYALLGCLLIGVSGSKLVKPEISLKGIOHVMQ
AGQILFKAGEAAHNSLPIIVSQEDKRSIIPPSACGRDNRQFCSLIIDIAQAH
TGLTYRSLTSTSKRKAESIIYFVSDAGSPFIEMHIDIPKLVEMEGROLIPCR
VTSNVIVLKPFQFDLPDQGRFIIDSRGFIIANATYKEGLINTEATVNGHLVQ
TNLTHPOINTILDOVIRPSPVRLHGOILVNLCTALIELATRVQMSWNPGRKAIK
ASIRQIDRSHSNVNHSHVSLKINNVESRDKGLYICVKSQSFQSTVSHVTEKGI
ISVHRKQPVQEIAGRRSIRLSMKVKAFFSPFVWLKDGSPAILKSARYLVHGISLI

```

```

IKDVTIEDAGYTILLGIKQSLFKNLTAIIVNVKPOIYKSVSLSPPLVPLGSR
QVLTICVYGIPIRPTIITWHPCHHNSKRDFTENESEFILDPSNLGNRIEISQ
RMVTEGNTKVTSLVADSQPGIYSCRAFNKIGTVERNIKFYVDVDPNGFHSLEK
MPAEGEDLKSLCVVNFYEDIDTWILLRTYVNRTHHNSISKQKMAITQDITSLNLVI
KNVLESDSTYACARNIYTCEDILRKTEVLVRDSEAPHLLQNLSDYEVISGSLTLD
CGARGVPAPQIWFKNHKIQOEPGIIILPGNSLTFLFERYEEDGVEGVCRAINQKA
VESAAYLVQIGISDKSNLELITCTVAAITLFWLLTFLIRKLRSSEVKIDYLSI
IMDPEVFLDEQERLPYDASKWEFARERLKLGLRGAGFAGVQVQASAFGKKSPIC
RIVAVKMLKEGATASEYKALMTFLKILTHIGHNLNVNLGACTQCGPLWIVFYCK
YGNLSNYLKSREDLFCINKDAALHMLAKESLEPLGQCGOKPRLDSSVSSSVSSPP
EDRSYDVEGEDYSEISKOPTMEDLISYFQVARGMEFLSRKCIHRDLAARNILL
SENKVKICDEGLARDIKNPQVRRGDTIRPLKWMAPESIFDKVITSKSDVSYGL
LWEIFSLGSPYFGVQDEDFCSRKEGMRMTPEYATPEIYQIMLDKWHKDPKPRR
FAELVEKLGDLQANVQOQDKYIPLNAILTRNSGFTYSTPTSEDLFKCGFADPHTH
SGSSDDRVYVNAFKFMSLERIKTFEELSPNSTSMFEDYQIDTSLILGSPILKRFTWE
TKPKASMKIDLRIAASKSKEAGLSDLPSPFCSSCGHIRPVQDSESELKESCCSPPP
DYSNVLYSSPPA"
316..412
/number=2
/evidence=experimental
413..542
/number=3
/evidence=experimental
543..757
/number=4
/evidence=experimental
758..930
/number=5
/evidence=experimental
931..1057
/number=6
/evidence=experimental
1058..1242
/number=7
/evidence=experimental
1243..1360
/number=8
/evidence=experimental
1351..1530
/number=9
/evidence=experimental
1531..1590
/number=10
/evidence=experimental
1591..1805
/number=11
/evidence=experimental
1806..1914
/number=12
/evidence=experimental
1907
/citation=[1]
/replace="t"
1915..2223
/number=13
/evidence=experimental
2224..2370
/number=14
/evidence=experimental
2371..2502
/number=15
/evidence=experimental
2503..2609
/number=16
/evidence=experimental
2610..2742
/number=17
/evidence=experimental
2743..2847
/number=18
/evidence=experimental
2848..2961

```

```

/number=19
/evidence=experimental
2562..3050
/number=20
/evidence=experimental
3051..3204
/number=21
/evidence=experimental
3205..3302
/number=22
/evidence=experimental
3303..3425
/number=23
/evidence=experimental
3426..3537
/number=24
/evidence=experimental
3538..3637
/number=25
/evidence=experimental
3638..3743
/number=26
/evidence=experimental
3744..3886
/number=27
/evidence=experimental
3792
/citation=[1]
/replace="a"
3886
/citation=[1]
/replace="c"
3887..3971
/number=28
/evidence=experimental
3972..4066
/number=29
/evidence=experimental
4067..5275
/number=30
/evidence=experimental
4115
/citation=[1]
/replace="g"
4254..5275
4464
/citation=[1]
/replace="c"
4517..4520
/citation=[1]
/replace="ccc"
5156
/citation=[1]
/replace="a"
5492
/citation=[1]
/replace="aaacct"
5954
/citation=[1]
/replace="s"
6275
polya_site
BASE COUNT 1665 a 1622 c 1584 g 1404 t
ORIGIN

```

```

Query Match      74.9%  Score 1574.5;  DB 10;  Length 6275;
Best Local Similarity 84.7%  Pred. No. 0;
Matches 1779;  Conservative 0;  Mismatches 319;  Indels 3;  Gaps 1;

QY  1  TATATCAGATGTCGAATGGTTTCATGTTGAAAATGCGGCGGANGA 60
    |||||||
Db  1905  TACGTACAGATGCGGATGGCTTTCACGTTTCCTGGAAGAGATCCAGCCGANGA 1964
    |||||||
QY  61  GAGGACCTGAACCTGCTTTCAGAGTTTACAACTTCTTAACAGAGAGCGTTACTTGGATT 120
    |||||||

```

```

Db  1965  GAGGACCTGAACCTGCTGTTGGTGAATAAATTCCTGACAGAGCAATTACCTGGATT 2024
    |||||||
QY  121  TACTGGGACAGTTAATAACAGAACTACCTACAGTAATTAGCAAGCAAAATGGCC 180
    |||||||
Db  2025  CTGCTACGGACAGTTAACAACAGAACCAATGACCATAGTATCAGCAAGCAAAATGGCC 2084
    |||||||
QY  181  ATCACTAAGGAGCAGCTCCATCACTCTTAATCTTACCATCATGAATGTTTCCCTGCAAGAT 240
    |||||||
Db  2085  ACCACCTCAAGATTACTCCATCACTCTGAACCTTGTCATCAAGAGCTGCTCTAGAGAC 2144
    |||||||
QY  241  TCAGGCACCTATGCTGAGAGCCAGGAATGTATACAGGGGAAATCTCCCAAG 300
    |||||||
Db  2145  TCGGCACCTATGCTGAGAGCCAGGAATATACAGGGGAAATCTCCCAAG 2204
    |||||||
QY  301  AAGAAATATCAATCAGAGATCAGGAAGCACCATACCTCTCTGCAAACTCAGTATCAC 360
    |||||||
Db  2205  ACAGAGTTCTGTTAGAGATTCGGAAGCCACACCTCTTCAAACTCAGTATCAC 2264
    |||||||
QY  361  ACAGTGGCCATCAGCAGTTCCACACTTTAGACTGTATGCTGTATGTTCCCGAGCCT 420
    |||||||
Db  2265  GAGTCTCCATCAGTGGCTCTACGACCTTAGACTGTAGAGTGTCCCGAGCCT 2324
    |||||||
QY  421  CAGATCAGTTGGTTTAAACAAACACACAAATACACAGAGCCTGGAATATTTAGGA 480
    |||||||
Db  2325  CAGATCAGTTGGTTTAAACAAACACACAAATACACAGAGCCTGGAATATTTAGGA 2384
    |||||||
QY  481  CCAGGAGCAGCAGCTGTTTATGAAGAGTCAGAGAGGAGTGAAGTGTCTCTATAGG 540
    |||||||
Db  2385  CCAGGAGCAGCAGCTGTTTATGAAGAGTCAGAGAGGAGTGAAGTGTCTCTATAGG 2444
    |||||||
QY  541  TCGAAAGCCACCAACAGAGGCTCTGTGGAAGTTCAGATACCTACTGTTCAAGGA 600
    |||||||
Db  2445  TCGGAGCCACCAACAGAGGCGCTGTGGAAGTTCAGATACCTACTGTTCAAGGA 2504
    |||||||
QY  601  ACTCGGACAGTCTAATCTGAGCTGATCACTTAACATGCACTGCTGCTGCTGCTG 660
    |||||||
Db  2505  ACTCGGACAGTCTAATCTGAGCTGATCACTTAACATGCACTGCTGCTGCTGCTG 2564
    |||||||
QY  661  CTCTCTGCTCTCTAATCAACCTCTTATCGGAAATGAAAAGTCTTCTTCTGAAATA 720
    |||||||
Db  2565  CTCTCTGCTCTCTAATCAACCTCTTATCGGAAATGAAAAGTCTTCTTCTGAAATA 2624
    |||||||
QY  721  AAGAGTACTACCTATCAATTAATGAGCCACAGATGAAGTCTCTTGTGATGAGCAGTGT 780
    |||||||
Db  2625  AAGAGTACTACCTATCAATTAATGAGCCACAGATGAAGTCTCTTGTGATGAGCAGTGT 2684
    |||||||
QY  781  GAGCGCTCTCTTATGATCCAGCAAGTGGAGTTTCCCGGGAGAGACTTAACTGGCC 840
    |||||||
Db  2685  GAGCGCTCTCTTATGATCCAGCAAGTGGAGTTTCCCGGGAGAGACTTAACTGGCC 2744
    |||||||
QY  841  AATACCTTGSNAGGGGCTTTTGGAAAGTGGTTCAAGCATCAGCAATTGGCAITTAAG 900
    |||||||
Db  2745  AATACCTTGSNAGGGGCTTTTGGAAAGTGGTTCAAGCATCAGCAATTGGCAITTAAG 2804
    |||||||
QY  901  AATACCTTACCTCCCGGAGCTGTGCTGTGAAATGCTGAAGAGGGGGCCACGGCCAGC 960
    |||||||
Db  2805  AATACCTTACCTCCCGGAGCTGTGCTGTGAAATGCTGAAGAGGGGGCCACGGCCAGC 2864
    |||||||
QY  961  GAGTACAAAGCTCTGATGACTGAGCTAAAATCTTGACCCACATTTGGCCACCTCTGAAC 1020
    |||||||
Db  2865  GAGTACAAAGCTCTGATGAGCTAAAATCTTGACCCACATTTGGCCACCTCTGAAC 2924
    |||||||
QY  1021  GTGTTAACTCTGAGGAGCTTGCACCAACAGAGAGGCGCTCTGATGGTATGTTGAA 1080
    |||||||
Db  2925  GTGTTAACTCTGAGGAGCTTGCACCAACAGAGAGGCGCTCTGATGGTATGTTGAA 2984
    |||||||
QY  1081  TACTGCAATATGAAATCTCTCAACTACCTCAAGAGCAACGCTGACTTATTTTCTC 1140
    |||||||
Db  2985  TACTGCAATATGAAATCTCTCAACTACCTCAAGAGCAACGCTGACTTATTTTCTC 3044
    |||||||
QY  1141  AAGAGATGCGAGCAGCTACATGAGCCTTAAGAGCAAAATGAGGAGCCAGCCTGGA 1200
    |||||||

```

Mon Dec 9 10:33:37 2002

Db 2045 AACAGGACCCAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGACAGCGCTGGAG 3104
 Qy 1201 CAAGGACGAGACAGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1260
 Db 3105 CAGGCGCAGAGAGCGCGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 3161
 Qy 1261 TTTCAGGAGAGATAAAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1320
 Db 3162 TTTCAGGAGAGATAAAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 3221
 Qy 1321 AAGGAGCGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 1380
 Db 3221 AAGGAGCGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 3281
 Qy 1381 GAGTTCCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 1440
 Db 3282 GAGTTCCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 3341
 Qy 1441 TCTGAGAGACAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 1500
 Db 3342 TCTGAGAGACAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 1560
 Qy 1501 CCGGAGTATGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1560
 Db 3402 CCGGAGTATGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1520
 Qy 1561 ATCTTGACAAATCTACAGACAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1520
 Db 3462 ATCTTGACAAATCTACAGACAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1520
 Qy 1621 GAAATCTTCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 1680
 Db 3522 GAGATCTTCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAGCGCTGGAG 1680
 Qy 1681 AGTCCCTGAGGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1740
 Db 3582 AGTCCCTGAGGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1740
 Qy 1741 CAGATATCTGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1800
 Db 3642 CAGATATCTGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1800
 Qy 1801 GTGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1860
 Db 3702 GTGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1860
 Qy 1861 CCAATCAATGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1920
 Db 3762 CCAATCAATGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1920
 Qy 1921 GAGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1980
 Db 3822 GAGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 1980
 Qy 1981 GTGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 2040
 Db 3882 GTGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 2040
 Qy 2041 CTTTACGAGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 2100
 Db 3942 CTTTACGAGAGAGAGATCTGAGCTTTCATATGAGCTCAGAAAGACAGAGCGCTGGAGCGCTGGAG 2100
 Qy 2101 G 2101
 Db 4002 G 4002
 RESULT 7
 LOCUS M. musculus Empr2 cDNA.
 DEFINITION X78568
 ACCESSION X78568.1 GI:510654
 VERSION

Empr2 gene: receptor kinase.
 Mus musculus.
 SOURCE
 ORGANISM
 Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
 1 (bases 1 to 5482)
 Chou, K., Wall, C., Hanratty, R. and Keller, G.
 Isolation of a gene encoding a novel receptor tyrosine kinase from differentiated embryonic stem cells
 Oncogene 9 (4), 1261-1266 (1994)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1..5482
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_type="embryonic stem cells"
 249..4241
 /gene="Empr2"
 249..4241
 /gene="Empr2"
 /codon_start=1
 /product="embryonic receptor kinase"
 /protein_id="CAA5531.1"
 /db_xref="GI:510655"
 /db_xref="MGI:95558"
 /translation="MVSQWDAVFLYFALVLCGLLLGSGGSKLVPELSLKGTHVMO
 AGQFLKCEGAHNSLPTVQEDKRLSTIPPSACGRONROFCSLTLDTAQANH
 TGLTICRLEPSTSKKKAESIVFSDAGSPFIEMHTDIPKLVHTEGROLICRY
 TSPNVTILAKFPFDLTDPDORITWDRGFIANATYKEIGLLCNCVANGHLYOTN
 YLTHQNTIILDVQIRPSPPLGLLGOILVNLCTATTELTVQWMSVNGKATKAS
 IQRLDRSHNNVHSHVILKINWESRDKGLYICRVKSGSPQSPSSVHVYKGLIK
 VKHRQVQETIAGRSYRLMKVAFSPETVWLKDGSPATLKSAFLVHGYSLIK
 TIEDAGDTILLGKQSLFKATLAILVNVKPOLYKESVSLPPLYSRQSM
 LICVYGPRTITLWHPDHHNHSKERYDECTEENESTLDPSSNLGNRIEYKMP
 TVIEGNTKTVLWADSGTPGYSCRAFNGKIGVERNIFVYVDPNGFHVLSKMP
 AEGDLKLSVWVKLYRDTITLLRTVNNRTHNSIKONGHSHSYSTILNIVKNP
 SLEDSTYACRNRIYGEDILRKTEVLVRDSEAPHLLOLSDYEVSIGSTLIDCOA
 RGVAPQITIKNNKHIQOEPGLILGPNSTLFIETVEDEGVYECSTNKGAVES
 AYLTVQGTSDKSNLELITVYCAALFLWLLTFLIRKLRSSEVKIDYLSIMD
 PVEVDEOCERLPYDASNEFARERKLKGLSLRGAGVQVQASVAGFKKSPICRY
 AVKMLKGAASEYKALMTELKILTHIGHLVNVLGACQKOGGPLMVIYVYKYN
 LSNYLSKRDLECLNDAALHMLKKELEPGLQEQPRUDSVSSSVTSSPFDNR
 SVSDVEGDEDESEISKOPLMEDLISYFQVARMFESSKCIHRDLAARNILSEN
 NVYKICDFGLARDIKNPDYVRGDTPLKWMAPESIFDKVYSTKSDVNSGVLWE
 IFSLGSPYGPVQWDEDECSRLKAGNMRTPEVATPEIYQIMLDCWHKDPKRFPAE
 LVEKLDLQANVQDQKDYIPLNAILIRNSNFTYSTPTFSEDRGKDGADPHFHSGS
 SDDVRYNAFKFMSLERIKTEELSPNSMPEDEYOLDSTLIGSLPKLKRFTETKP
 KASKMIDLRITAFSPKEAGLSLPRPFCSCGHIRPVODDESELGKESCCSPPDYN
 SVLYSSPPA"
 BASE COUNT 1464 a 1446 c 1369 g 1203 t
 CIGIN
 Query Match 73.0%; Score 1534.2; DB 10; Length 5482;
 Best Local Similarity 84.1%; Pred. No. 0;
 Matches 1767; Conservative 0; Mismatches 328; Indels 6; Gaps 3;
 Qy 1 TATATCAGATGTCGCAAAATGGTTCATGTAACTTGGAAAAATGCGAGCGAAGGA 60
 Db 1896 TACGTACAGATGTCGCAAAATGGTTCATGTAACTTGGAAAAATGCGAGCGAAGGA 60
 Qy 61 GAGGACCTGAAACTGTCTTCACAGTTAACAAGTTCCTTATACAGACCTTACTTGATT 120
 Db 1956 GAGGACCTGAAACTGTCTTCACAGTTAACAAGTTCCTTATACAGACCTTACTTGATT 120
 Qy 121 TTACTCGGACAGTATATACAGAAATCCACTACACTTATACAGCAAAAATGGCC 180

Db	2016	CTGCTACGACAGTHTAACACAGAACCCATGCAATAGTATCAGCAAGC--AAATGCCC	2073
Qy	181	ATCATTAGGAGCACTCCATCACTCTTAATCHAACCAATGAAATGTTTCCCTCGAAT	240
Db	2074	ACCACTCAAG-TTACTGCATCTGTGAACCTGTGTAACAGAGCTGTCTTGAAGAC	2132
Qy	241	TCAGCACCTATGCTTCGAGAGCCAGGAATGTATACAGGGGAGAAATCTTCCCAAG	300
Db	2133	TCGGCACCTATGCTGTCAGAGCCAGGACATATACAGAGGAGACATCTTCGGAAG	2192
Qy	301	AAAGAAATTAACAATCAGAGTACAGAGCACCAATACCTCTTCGGAACCTCACTGATAC	360
Db	2193	ACAGAAGTCTCGTTAGAGATTCGGAAGCCACACCTGCTTCAAACTCTAGTAC	2252
Qy	361	ACAGTGGCAATCAGCAGTTCACCACTTTAGAGCTCAAGCTAATGCTGCTCCCGAGCT	420
Db	2253	GAGGCTCCATCAGTGGCTCTAGACCTTAGACTGTCAAGCTAGAGTGTCCCGGCT	2312
Qy	421	CAGATCACTTGGTTTAAAAACACCACAAATATACAGAGCCCTGGAAATATTTAGGA	480
Db	2313	CAGATCACTTGGTTTAAAAACACCACAAATATACAGAGCCCTGGAAATATTTAGGA	2372
Qy	481	CCAGAGACAGCTGCTGTTTATGNAGAGTACAGAGAGGATGAGGTGTCTATAC	540
Db	2373	CCAGAAACAGCAGCTGTTTATGAAAGAGTACAGAGGAGGATGAGGTGTCTATAC	2432
Qy	541	TGCAAGGCATACCCAGAGGGCTCTGCGAATTCAGCATACCTCTGATGAGT	600
Db	2433	TGCGATCCACCAACAGAGGGGGCTGGAAAGCCAGCTACCTACCGTGAAGA	2492
Qy	601	ACCTCGCAAGCTCTAATCTGGAGCTGATCACTCTAATATGCACTGTGTGCTGCAC	660
Db	2493	ACCTCAGACAGTCAACCTGGAGCTGATCAGCTCAGCTACATGCGGCTGGGCT	2552
Qy	661	CTCTTCTGGCTCTAATAGCCCTCTATCCGAAAAGAAAGTCTTCTTGAATA	720
Db	2553	CTCTTCTGGCTCTTACTCTCTCAAGAAATGCAAGCGTCTCTTCGGAAT	2612
Qy	721	AAGCTGACTACCTATCAATTAATGAGCCAGATGAGTTCCTTTGGATGAGAGTGT	780
Db	2613	AAGACAGACTTCCCTGTAATCAATATGACCAGATGAGTTCCTTGGATGAGAGTGT	2672
Qy	781	GAGCGCTCCCTTATCATGCCAGCAAGTGGAGTTTGCCTGGGAGAGACTTAACTG	840
Db	2673	GACCGCTGCCCTATGATGCCAGCACTGGAGTTTGCAGGAGAGACTGAATAGC	2732
Qy	841	AAATCACTTGAAGAGGGCTTTGAAAGTGTTCAGCATCAGCATTTGGCATTAAG	900
Db	2733	AAATCGCTCGGAAGGGCTTTTGGAAAGTGTTCAGCATCAGCATTTGGCATTAAG	2792
Qy	901	AAATCACTTACGTGCCGGACTGTGCTGTGAAATGCTGAAGAGGGGGCCAGGCCAG	960
Db	2793	AAATCACTTACGTGCCGGACTGTGCTGTGAAATGCTGAAGAGGGGGCCAGGCCAG	2852
Qy	961	CAGTACAAGCTCTCATCACTTCACTTAAATCTTGACCCACTTGGCCACCATGTAAC	1020
Db	2853	GAGTACAAGCTCTCATCACTTCACTTAAATCTTGACCCACTTGGCCACCATGTAAC	2912
Qy	1021	GTGTTTAACTCTTGGAGCTGCACCAAGCAAGAGGGCTCTGATGTGATGTTGA	1080
Db	2913	GTGTTTAACTCTTGGAGCTGCACCAAGCAAGAGGGCTCTGATGTGATGTTGA	2972
Qy	1081	TACTGCAATATGAAATCTCCAACTACCTCAAGAGCAAGCTGACTTAATTTCTC	1140
Db	2973	TACTGCAATATGAAATCTCCAACTACCTCAAGAGCAAGCTGACTTAATTTCTC	3032
Qy	1141	AACAAGATGACCATACATGAGCGCTTAGAAAGAAAATGGACCGAGCGCTGGAA	1200
Db	3033	AACAAGATGACCGCTTGCATGAGCTCAAGAAAGAGCGCTTGGAAACAGCGCTGAG	3082
Qy	1201	CAAGGCAAGAACCAAGCTAGATAGAGTCAACAGAGCAAGCTTTGGAGTTCGGC	1260
Db	3093	CAGGGCAAGAGCGCGCTAGACAGTGT---CAGAGCTCAAGTGTACCAAGTCCAG	3120

CF	1261	TTTCAGGAGAGATAAAAGTCTCAGTGTGTTGAGGAGAGGAGGATTCTGACGGTTTCTTC	1320
DB	3150	ITCCCTGAGAGCCGAGCGTGAGCGATGTGAGGAGGACGAGGATTCAGTGTAGATCTCC	3209
QY	1321	AAGGAGCTCATCTATCGGAAGATCTGATTTCTTTACAGTCTTTCAGTGCCGAGAGCA13	1380
DB	3210	AAGCAGCGCTCACCATGGAAGACCTGATTTCTACAGTTTCCAAGTGGCCGAGAGCA13	3269
QY	1381	GAGTCTGCTTCTCCAGAAGTGCAATTCGGGACCTGGGAGCGAGGAGGAGACATCTCTTTA	1440
DB	3270	GAGTTCTCTGCTCCAGAAAGGAGATTCATCGGACCTGGGAGCGAGGAGACATCTCTTTA	3329
QY	1441	TTTGAGAAACAGCTGGTTCAGATTTGTGATTTTGGCTTCCCGGGGATATTTATAAGAAC	1500
DB	3330	TTTGAGAAACAGCTGGTTCAGATTTGTGATTTTGGCTTCCCGGGGATATTTATAAGAAC	3389
QY	1501	CCGGATTTGTGAAAAGGAGATCTCGATCTCTCTGAAATGGATGCTCCCGAATCT	1560
DB	3390	CTGGATTTGTGAAAAGGAGATCTCGATCTCTCTGAAATGGATGCTCCCGAATCT	3449
QY	1561	ATCTTGGAGAAATCTACAGCAACAGGAGGAGCTGTGCTTTACGAGTATTTGCTGAG	1620
DB	3450	AATCTTGGAGAAATCTACAGCAACAGGAGGAGCTGTGCTTTACGAGTATTTGCTGAG	3509
QY	1621	GAATCTCTCTTAGGTGGGTCTCCATACCCAGAGTACAAATGGATGAGGATTTGCTGAG	1680
DB	3510	GAGATCTCTCTTAGGTGGGTCTCCATACCCAGAGTACAAATGGATGAGGATTTGCTGAG	3569
QY	1681	AGTGGCTGTGAGGAGGATGAGGATGAGGATCTCTGAGTACTTACTCTGAAATCTAT	1740
DB	3570	AGCCGCTGTGAGGAGGATGAGGATGAGGATCTCTGAGTACTTACTCTGAAATCTAT	3629
QY	1741	CAGATCTGTGAGTCTCTGAGGAGGATGAGGATGAGGATCTCTGAGTACTTACTCTGAAATCTAT	1800
DB	3630	CAGATCTGTGAGTCTCTGAGGAGGATGAGGATGAGGATCTCTGAGTACTTACTCTGAAATCTAT	3689
QY	1801	GTGAAAAAATCGTGTGATTTGCTTCAAGCAAAATGTACACAGGATGTGAAGACTACATC	1860
DB	3690	GTGAAAAAATCGTGTGATTTGCTTCAAGCAAAATGTACACAGGATGTGAAGACTACATC	3749
QY	1861	CCAATCAATGCCATACCTACAGAGAAATAGTGGTTTACATCTCAACTCTGCTCTCTCT	1920
DB	3750	CCAATCAATGCCATACCTACAGAGAAATAGTGGTTTACATCTCAACTCTGCTCTCTCT	3809
QY	1921	GAGGACTCTTTAGGAGGATTTTACGCTCCGAGTTTAAATTCAGAACTCTGATGAT	1980
DB	3810	GAGGAGCGGCGAGGAGGCTTTGAGATCCCATTTTAAATTCAGAACTCTGATGAT	3869
QY	1981	GTGAGATGATGATCTTTTCAAGTTTCAAGCTTGAAGGATGGAAGGATCAAACTTTGAAGA	2040
DB	3870	GTGAGATGATGATCTTTTCAAGTTTCAAGCTTGAAGGATGGAAGGATCAAACTTTGAAGA	3929
QY	2041	CTTTACCGAAATCCCACTCTCATTTTATGATCTACACGGGCGAGCAGCTCTGTG	2100
DB	3930	CTTTACCGAAATCCCACTCTCATTTTATGATCTACACGGGCGAGCAGCTCTGTG	3980
QY	2101	G 2101	
DB	3990	G 3990	

RESULT 8
AB065372
LOCUS

DEFINITION
 Gallus gallus mRNA
ACCESSION
 AB065372
VERSION
 AB065372.1
KEYWORDS
 Gallus gallus tissue lib.whole embryo mRNA.

4272 bp
 linear
 VRT 28-MAY-2002
 complete cds.
 embryonal growth factor

PEPCLI 8	AB065372	LOCUS	AB065372	DEFINITION	Gallus gallus 4272 bp mRNA receptor-1, complete cds.	ACCESSION	AB065372	VERSION	AB065372.1	KEYWORDS	GI:18652860	SOURCE	Gallus gallus tissue lib:whole embryo mRNA.	ORGANISM	Gallus gallus	EXAMINOTA	Vertebrata; Chordata; Craniata; Vertebrata; Euteleostomi;
----------	----------	-------	----------	------------	--	-----------	----------	---------	------------	----------	-------------	--------	---	----------	---------------	-----------	---


```

* 129384 129483: gap of 100 bp
* 129484 131147: contig of 2264 bp in length
* 131148 131847: gap of 100 bp
* 131848 132316: contig of 469 bp in length
* 132317 132416: gap of 100 bp
* 132417 134455: contig of 2039 bp in length
* 134456 134555: gap of 100 bp
* 134556 135527: contig of 972 bp in length
* 135228 135627: gap of 100 bp
* 135628 189051: contig of 53424 bp in length
* 189052 189151: gap of 100 bp
* 189152 189476: contig of 325 bp in length
* 189477 189576: gap of 100 bp
* 189577 191375: contig of 1799 bp in length
* 191376 191475: gap of 100 bp
* 191476 201473: contig of 999 bp in length
* 201474 201573: gap of 100 bp
* 201574 202307: contig of 734 bp in length
* 202308 202407: gap of 100 bp
* 202408 204878: contig of 2471 bp in length
* 204879 204978: gap of 100 bp
* 204979 213331: contig of 8553 bp in length
* 213332 213631: gap of 100 bp
* 213632 218109: contig of 4478 bp in length
* 218110 218209: gap of 100 bp
* 218210 219800: contig of 1591 bp in length
* 219801 219900: gap of 100 bp
* 219901 220965: contig of 1065 bp in length.
FEATURES
  source
    Location/Qualifiers
      1. .220965
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="3"
        /map="3p21.3"
        /clone="RP6-32g23"
BASE COUNT  62577 a 44674 c 45661 g 65453 t 2600 others
ORIGIN
  Query Match      45.5%; Score 956.4; DB 2: Length 220965;
  Best Local Similarity 94.6%; Pred. No. 6.4e-268;
  Matches 990; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 1056 AGGCGCTCTGATGCTGATTTGAACTGCAAAATGGAATCTCTCCAACTACCTCAA 1115
Db 25255 AGGCGCTCTGATGCTGATTTGAAATGGAATGGAATCTCTCCAACTACCTCAA 25196
QY 1116 GAGCAACGTCACCTATTTTCTCAACAGGATGACAGCTACACATGGAGCGCTAGAA 1175
25195 GAGCAATATGACTTATTTTCTGCAACAGGATGTCACATGAGCGCTAGAA 25136
QY 1176 AGAAAAATGAGCCAGGCTCGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1235
Db 25135 AGAAAAATGAGCCAGGCTCGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 25075
QY 1236 CAGCGAAGTTTTCGAGCTCGGCTTTCAGAGAGATAAAAGTGTAGTGTAGGAA 1295
Db 25075 CAGCGAAGTTTTCGAGCTCGGCTTTCAGAGAGATAAAAGTGTAGTGTAGGAA 25016
QY 1296 AGAGCAGGATCTCAGCGGTTCTCAGAGAGCCCATCACTATGAGATCTCATTTCTTA 1355
Db 25015 AGAGGAGATCTCAGCGGTTCTCAGAGAGCCCATCACTATGAGATCTCATTTCTTA 24956
QY 1356 CAGTTTTCAGATGCGCAGAGGATGAGTTCTCTGCTTCCAGAAAGTCAATCATCGGGA 1415
Db 24955 CAGTTTTCAGATGCGCAGAGGATGAGTTCTCTGCTTCCAGAAAGTCAATCATCGGGA 24896
QY 1416 CTGCGCAGGAGAAATCTTTTATCTAGAGAAACAGTGTTGAGAGATTTGATTTGG 1475
Db 24895 CTGCGCAGGAGAAATCTTTTATCTAGAGAAACAGTGTTGAGAGATTTGATTTGG 24836
QY 1475 CTTTCCCGCGGATATTTATAGAACCCCGATTTATCTGAGAAAGAGAGATCTGACATCTC 1535
Db 24835 CTTTCCCGCGGATATTTATAGAACCCCGATTTATCTGAGAAAGAGAGATCTGACATCTC 24776

```

```

QY 1536 ICTGAAATGGATGCTCCGGAATCTATCTTTGACAAAATCTACAGCACCAGAGCGAGT 1595
Db 24775 ICTGAAATGGATGCTCCGGAATCTATCTTTGACAAAATCTACAGCACCAGAGCGAGT 24716
QY 1596 GTGTCCTTACGGAGATGCTGCGGAAATCTCTCTAGTGGGTCTCCATACCAGG 1655
Db 24715 GTGTCCTTACGGAGATGCTGCGGAAATCTCTCTAGTGGGTCTCCATACCAGG 24656
QY 1656 AGTCAAAATGGATGAGGAGCTTTTTCAGTTCGCTGAGGAAAGCATGAGGATGAGAGCTCC 1715
Db 24655 AGTCAAAATGGATGAGGAGCTTTTTCAGTTCGCTGAGGAAAGCATGAGGATGAGAGCTCC 24596
QY 1716 TGATGATCTACTCTCTGAAATCTATCAGATCATCTGGAGCTGCTGCGACAGACCCAAA 1775
Db 24595 TGATGATCTACTCTCTGAAATCTATCAGATCATCTGGAGCTGCTGCGACAGACCCAAA 24536
QY 1776 AGAAGGCGCAAGATTTGCAAGACTTGTGNAAAACTAGGTGATTTGCTTCAAGCAATGT 1835
Db 24535 AGAAGGCGCAAGATTTGCAAGACTTGTGNAAAACTAGGTGATTTGCTTCAAGCAATGT 24475
QY 1836 ACAACAGATGTTAAAGACTTACATCCCAATCAATGCCATCTACAGAGAAATAGTGGGT 1895
Db 24475 ACAACAGATGTTAAAGACTTACATCCCAATCAATGCCATCTACAGAGAAATAGTGGGT 24416
QY 1895 TACATATCAACTCTGCTGCTTCTGAGAGCTTCTCAAGAAAGTATTCAGCTCCGAA 1955
Db 24415 TACATATCAACTCTGCTGCTTCTGAGAGCTTCTCAAGAAAGTATTCAGCTCCGAA 24356
QY 1956 GTTAAATTCAGGAAGCTCTGATGATGTCAGATATGTAATGCTTCAAGTTCATGAGCCT 2015
Db 24355 GTTAAATTCAGGAAGCTCTGATGATGTCAGATATGTAATGCTTCAAGTTCATGAGCCT 24296
QY 2016 GGAAGAAATCAAAACCTTTTGAAGAACTTTTACCAGATGCCCAATGCCATGTTTATGATGACTA 2075
Db 24295 GGAAGAAATCAAAACCTTTTGAAGAACTTTTACCAGATGCCCAATGCCATGTTTATGATGACTA 24236
QY 2075 CCAGGGGACAGGAGGAGCTCTGTGG 2101
Db 24235 CCAGGGGACAGGAGGAGCTCTGTGG 24210
LOCUS AC104439 197279 bp DNA linear PRI 20-JUN-2002
AC104439/c
DEFINITION Homo sapiens chromosome 3 clone RP11-793E15, complete sequence.
ACCESSION AC104439 AC024739
VERSION AC104439.2 GI:21490240
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 197279)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197279)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 197279)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Jun 20, 2002 this sequence version replaced gi:1748621.
----- Genome Center

```

Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgchgs@u.washington.edu
 Drafting Center: WUGSC
 ----- Project Information
 Center project name: chr-3
 Center clone name: RP11-793E15 (bc0564)
 ----- Summary Statistics

Sequencing vector: unknown; 52% of reads
 Sequencing vector: plasmid; 108752: 48% of reads
 Chemistry: Dye-terminator EI; 94% of reads
 Chemistry: Dye-terminator Big Dye; 6% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 197168 bases at least Q40
 Consensus quality: 197255 bases at least Q30
 Consensus quality: 197275 bases at least Q20
 Insert size: 197279; sum-of-contigs
 Quality coverage: 8.2x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP11-91E8 (UWGC:bc0216) AC026349
 3': CTD-2563A18 (UWGC:bc0730)

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
 all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

HindIII			BglII			EcoRI		
SeqDerMap	FrgrPrint	SeqDerMap	FrgrPrint	SeqDerMap	FrgrPrint	SeqDerMap	FrgrPrint	SeqDerMap
2687	2617	8949	8586	8596	8661			
6382	6410	2067	2160	6	<800			
512	<800	7846	7940	2742	2803			
449	<800	3734	3895	5376	5324			
2602	2763	1334	1301	1433	1478			
2590	2617	2287	2309	823	835			
8313	8291	1814	1918	1962	2002			

1711	1683	691	<800	2900	2954
9821	9472	5477	5348	1484	1478
516	<800	305	<800	1005	995
8587	8291	25245	25541	1181	1171
7446	7581	3988	4121	18560	19002
2088	2075	1633	1598	3503	3579
2509	2617	631	<800	4943	5076
3519	3501	90	<800	3239	3241
26	<800	402	<800	953	995
925	933	3350	3490	1621	1615
98	<800	4577	4515	6827	6900
6409	6410	1229	1183	79	<800
1383	1376	4221	4331	16418	16263
16048	15829	3980	4121	3361	3241
1876	1877	2548	2497	872	881
4169	4068	674	<800	4220	4193
1674	1683	2332	2309	2848	2803
52	<800	11445	11045	2283	2318
5227	5082	5692	5632	3211	3241
15464	15829	1385	1414	4615	4632
14333	14296	608	<800	6823	6900
1026	1026	3930	3895	886	881
645	<800	1465	1414	2032	2002
5621	6410	2382	2497	8680	8661
8413	8291	747	<800	5163	5076
1401	1376	1951	2012	1547	1478
1512	1471	642	<800	7058	6900
5801	5644	21060	21003	1647	1615
3256	3278	3700	3666	2753	2803
3448	3501	1321	1301	162	<800
2738	2763	10705	10502	1078	1093
237	<800	3128	3317	2156	2173
6715	6823	183	<800	5077	5076
201	<800	3255	3490	154	<800
1002	1026	5693	5632	1571	1478
221	<800	3380	3317	219	<800
2854	2899	237	<800	520	<800

```
----- 7317 431 ----- 1540 1478
----- 1145 2152 ----- 1479 1478
----- 8291 249 ----- 1513 1615
----- 1877 2599 ----- 1399 1478
----- 3847 695 ----- 9553 9383
----- 5644 41 ----- 155 <800
----- 1711 1683 ----- 13615 13115
----- 852 ----- 4731 4632
----- 2389 ----- 5159 5076
----- 178 ----- 4503 4430
----- 6589 ----- 6607 6900
----- 3836 ----- 2766 2803
-----

Query Match 45.4%; Score 953.2; DB 9; Length 197279;
Best Local Similarity 94.5%; Pred. No. 5.4e-267;
Matches 988; Conservative 0; Mismatches 58; Indels 0; Gaps 0:

QY 1056 AGGGCTCTGATGTTGTTGAATACCTGCAATATGGAATCTCTCCAACTACCTCAA 1115
DB 68708 AGGGCTCTGATGTTGTTGAATATGCAATATGGAATCTCTCCAACTACCTCAA 68649
QY 1116 GACCAAGCTGACATTTTCTCAACAGATGACGACCTACATGAGCCCTAAGAA 1175
DB 68648 GAGCAATATGACATTTTCTCAACAGATGACGACCTACATGAGCCCTAAGAA 68589
QY 1176 AGAAAAATGGAGCGCTGGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1235
DB 68588 AGAAAAATGGAGCGCTGGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 68529
QY 1236 CAGCGAAGCTTTTCGAGCTCCGGCTTTTCAAGCAAGATAAAGTCTGAGTGTG 1295
DB 68528 CAGCGAAGCTTTTCGAGCTCCAAAGTTTCAGCAAGATAAAGTCTGAGTGTG 68469
QY 1296 AGAGGAGATCTCACGTTTCTACAGAGCCCATCACTATGAGATCTCAATCTTA 1355
DB 68468 AGAGGAGATCTCACGTTTCTACAGAGCCCATCACTATGAGATCTCAATCTTA 68409
QY 1356 CAGTTTCAAGTGGCCAGAGCAAGGATTCCTGTCTTCCAGAAAGTGCATTCAGGGA 1415
DB 68408 CAGTTTCAAGTGGCCAGAGCAAGGATTCCTGTCTTCCAGAAAGTGCATTCAGGGA 68349
QY 1416 CTGGCAGGAGACATCTTATCTGAGACAGCTGGTGAAGATTTGTAATTTGG 1475
DB 68348 CTGGCAGGAGACATCTTATCTGAGACAGCTGGTGAAGATTTGTAATTTGG 68289
QY 1476 CTTTCCCGGGATATTTTAAAGAACCCCGATTTCTGAGAAAAGGAGATCTGACATCC 1535
DB 68288 CTTTCCCGGGATATTTTAAAGAACCCCGATTTCTGAGAAAAGGAGATCTGACATCC 68229
QY 1536 TGTAAATGGATGGCTCCCGGATCTATCTTTGACAAAATCTACAGCAAGAGCGAGT 1595
DB 68228 TGTAAATGGATGGCTCCCGGATCTATCTTTGACAAAATCTACAGCAAGAGCGAGT 68169
QY 1596 GTGGTCTTACGAGTATCTGTGGGAATCTTCTTGGTGGTCTCCATCCCGAGG 1655
DB 68168 GTGGTCTTACGAGTATCTGTGGGAATCTTCTTGGTGGTCTCCATCCCGAGG 68109
QY 1656 AGTACAAATGGATGAGGACATTTTGCAGTGGCTGAGGAGGAGATGAGGATGAGGCTCC 1715
DB 68108 AGTACAAATGGATGAGGACATTTTGCAGTGGCTGAGGAGGAGATGAGGATGAGGCTCC 68049
```

```
QY 1716 TGAGTACTTACTCTCTGAAATCTATCAGATCATCTGGACTGTGGCAGAGACCCAAA 1775
DB 68048 TGAGTACTTACTCTCTGAAATCTATCAGATCATCTGGACTGTGGCAGACCAAGACCCAAA 67989
QY 1776 AGAAGGCCAGATTTTGCAGAACTTGTGAAAACCTAGGTGATTTCTCAAGCAAAATGT 1835
DB 67988 AGAAGGCCAGATTTTGCAGAACTTGTGAAAACCTAGGTGATTTCTCAAGCAAAATGT 67929
QY 1836 ACAACAGGATGTTAAAGACTACATCCCAATCAATCCCATCTGACAGGAAATAGTGGGT 1895
DB 67928 ATPACAGATGTTAAAGACTACATCCCAATCAATCCCATCTGACAGGAAATAGTGGGT 67869
QY 1896 TACATCTCAACTCTCTGCTCTCTCTGAGGACTTTTCAAGGAAAGTATTTTCAAGTCCGAA 1955
DB 67868 TACATCTCAACTCTCTGCTCTCTCTGAGGACTTTTCAAGGAAAGTATTTTCAAGTCCGAA 67809
QY 1955 GTTAAATCAGGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2015
DB 67808 GTTAAATCAGGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 67749
QY 2016 GGAAGAATCAAAACCTTTTGAAGAATTTTACCAGATGCCACCTCCATGTTTCATGACTA 2075
DB 67748 GGAAGAATCAAAACCTTTTGAAGAATTTTACCAGATGCCACCTCCATGTTTCATGACTA 67689
QY 2076 CCAGGGCCAGCAGCAGCTCTCTGCTGG 2101
DB 67688 CCAGGGCCAGCAGCAGCTCTCTGCTGG 67663

RESULT 11
AX481480
LOCUS AX481480 4044 bp DNA linear PAT 15-AUG-2002
DEFINITION Sequence 94 from Patent WO02055693.
ACCESSION AX481480
VERSION AX481480.1 GI:22316394
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kreutzer,R., Limmer,S., Rost,S. and Hadwiger,P.
TITLE Method for inhibiting the expression of a target gene
JOURNAL Patent: WO 02055693-A 94 18-JUL-2002;
Ribopharma AG (DE)
FEATURES
Source Location/Qualifiers
1..4044
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1153 a 885 c 1021 g 975 t
ORIGIN

Query Match 34.0%; Score 713.8; DB 6; Length 4044;
Best Local Similarity 65.5%; Pred. No. 3.8e-197;
Matches 1077; Conservative 0; Mismatches 562; Indels 5; Gaps 2:

QY 218 TCATGAATCTTCCCTCCAGATTCAGGACCTATGCTGCGAGAGCCAGGAATGATACA 277
DB 1886 TTAGATATGATCTCTGAGGACCAAGGAGACTATGCTGCTTCTCAAGACAGGAAGA 1945
QY 278 CAGGGAAGAAATCTCTCCAGAGAAATTTACAATCAGAGATCAGGAAGCACCATACC 337
DB 1945 CCAAGAAAGACATTTGGTGGTCCAGGCGCTACAGTCTTAGAGCGGTGGCCACCCAGGA 2005
QY 338 TCTGCGAAACTCAGTATATCAGAGTGGCCATCAGAGTTCACCACTTTTAGACTGTC 397
DB 2006 TCACAGGAACCTTGAGATTCAGACAGCAAGTATGGGAAAGCAATCGAAGTCTCATGA 2065
QY 398 ATGCTAATCTGTCCTCCGAGCCTTCAATCCTTGGTTTAAACACACCAACATACAA 457
DB 2065 CGGCATCTGGGATCCCTCCACAGATCATGTGTTTAAAGATAATGAGACCTTGTAG 2125
```


Db 2246 AGCATTTTCATTAATGAGGTCGCCAGAAAGCAACTTGGAAATCATTAATCTAG 2305
Qy 638 CATGCACCTGTGTGCTGGACCTCTCTTCTGGCTCTCTATTAACTCTCTTATCCGAAAAA 697
Db 2306 TAGGCACGCGGTGATGGCAATGCTTCTGGCTCTCTTCTGGCTCTCTTATCCGAAAAA 2365
Qy 698 TGAAGAGTCT---TCTCTTGAATTAAGACTGATCTATCAATTAATTAAGGACCCAG 754
Db 2366 TTAAGCGGCGCAATGGAGGGGAACCTGAAGACAGGATCTCTCAATGCTGATCCAG 2425
Qy 755 ATGAAGTCTCTTGTGATGAGCTGTGAGCGGCTCTCTTATGATGCGCAGCAAGTGGAGT 814
Db 2426 ATGACTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2485
Qy 815 TTGCCCGGAGAGACTTAAGCTGGGCAAAATCACITGGAGAGGCGCTTTTGGAAAGTGG 874
Db 2486 TCCCCAGAGACCGGTGAAGTAGTAAGGCTCTTGGCGGTGGTGGCTTTGGCAAGTGA 2545
Qy 875 TTCAAGCATCAGCATTTGGCAATTAAGATACCTACCTACCTGGGAGCTGTGGCTGTGAAA 934
Db 2546 TTGAACAGATGCTTTGGCAATTAAGATACCTACCTACCTGGGAGCTGTGGCTGTGAAA 2605
Qy 935 TGTCTGAAGAGGCGCCAGCGGAGTACAAAGCTCTGATGACTGAGCTAANAATCT 994
Db 2606 TGTGAAGAGGAGCAACACAGAGTGAAGATGAGCTCTGATGACTGAGCTAANAATCT 2665
Qy 995 TGACCCACATTTGGCACCACATCTGAGCTGTTAACTCTGGGAGCTGTGGCTGTGAAA 1054
Db 2666 TCAATCATATTTGGCACCACATCTGAGCTGTTAACTCTGGGAGCTGTGGCTGTGAAA 2725
Qy 1055 GAGGCGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1114
Db 2726 GAGGCGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2785
Qy 1115 AGGCAAACTGCTATTTTCTCAAGAGGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1174
Db 2786 GGAGCAAGAGAAATGAATTTGCTCCCTACAAAGCAAGGCGGAGCTGCTGAGGAG 2845
Qy 1175 AAGAAATAATGGAGCGGCTGGACAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1234
Db 2846 AAGACTAGCTGTGGAGCAATCCCTG---TGATCTGAAACGCGCTTGGAGAGCTACCA 2902
Qy 1235 GCAGCAAAAGCTTTGCGAGCTCCGCTTTAGGAAGATTAAGCTGAGCTGAGCTGAGCT 1294
Db 2903 GTAGCCAGAGCTACAGGAGCTGTGATTTGAGGAGAGAGTCCCTGAGTGTAGAG 2962
Qy 1295 AAGAGAGATTTCTAGCGTTTCTCAAGAGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1354
Db 2963 AAGAGAGATTTCTAGCGTTTCTCAAGAGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCT 3022
Qy 1355 ACAGTTTTCAGTGGCCAGAGGAGTGGAGTCTGCTTCCAGAGAGTGGAGTGGAGTGGAG 1414
Db 3023 ACAGTTTTCAGTGGCCAGAGGAGTGGAGTCTGCTTCCAGAGAGTGGAGTGGAGTGGAG 3082
Qy 1415 ACTGCGAGCGAGAAATCTTTTATCTGAGCAAGCTGGTGAAGATTTGTGATTTG 1474
Db 3083 ACTGCGCGCAGAAATCTTTTATCTGAGCAAGCTGGTGAAGATTTGTGATTTG 3142
Qy 1475 GCCTTGGCGGAGATTTTATGAGAGCGGCTTATGAGAGAGAGAGAGTCTGAGCTT 1534
Db 3143 GCCTTGGCGGAGATTTTATGAGAGCGGCTTATGAGAGAGAGAGAGTCTGAGCTT 3202
Qy 1535 CTCTGAATGAGTGGCTCCGAGATCTATCTTGAAGAAATCTAGAGTACAGAGAGAGAG 1594
Db 3203 CTCTGAATGAGTGGCTCCGAGATCTATCTTGAAGAAATCTAGAGTACAGAGAGAGAG 3262
Qy 1595 TGTGCTTACGAGATTTGCTGAGAGATCTTCTCTGAGTGGTGTCTGCAATCCAG 1654
Db 3263 TGTGCTTACGAGATTTGCTGAGAGATCTTCTCTGAGTGGTGTCTGCAATCCAG 3322
Qy 1655 GAGTCAAAATGATGAGAGCTTTTGTAGTGGCTGAGGAGAGAGAGAGAGAGAGAGAG 1714

Db 3323 GGTAAGAGATGATGAGAAATTTTGTAGGCGATTTGAAAGAGAAAGTAAAGTAGAGGCC 3382
Qy 1715 CTGAGTACTTACTCTCTGAAATCTATCATGATCATCTGAGTCTGTGCGACAGACCCAA 1774
Db 3383 CTGATATATCTACACCAAGAAATGACACAGACCATCTGAGTCTGTGCGACGGGAGCCCA 3442
Qy 1775 AGAAGGCGCAAGATTTGCGAGACTTTGCGAAGAACTAGTGTGATTTGCTTCAAGCAATG 1834
Db 3443 CTGAGAGACCCAGCTTTTCAAGTGTGTGGAACATTTGGGAATCTCTTGAAGTAAATG 3502
Qy 1835 TACAGCAGATGTAAGACTACAT 1859
Db 3503 CTGAGCAGATGTAAGACTACAT 3527

PFRESULT 13
AP201382
LOCUS AR201382 4071 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6359115.
ACCESSION AR201382
VERSION AR201382.1 GI:20252270
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4071)
AUTHORS Rendall, R.L., Mao, X., Thomas, K.A. and Tebben, A.
TITLE Human receptor tyrosine kinase, KDR
JOURNAL Patent: US 6359115-A 19-MAR-2002;
FEATURES Location/Qualifiers
Source 1..4071
BASE COUNT 1169 a 894 c 1026 g 982 t
CPIDIN

Query Match 34.0%; Score 713.8; DB 6; Length 4071;
Best Local Similarity 65.5%; Pred No. 3.8e-197;
Matches 1077; Conservative 0; Mismatches 582; Indels 6; Gaps 2;

Qy 218 TCAAGATGTTTCCCTGCAAGATTCAGGACCTATGCTCTGAGAGCGGAGGAGTATACA 277
Db 1886 ITAAGATGCTATCTTGGAGGAGGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1945
Qy 278 CAGGGAAGAAATCTCCAGAGAGAAATTAATACATCAGAGATCAGAGAGAGAGAGAG 337
Db 1846 CCAAGAAAGAGATGCTGCTGAGGAGCTCACAGTCTTAGAGCGTGTGGACCCAGGA 2005
Qy 338 TCTGCGAAACCTCAGTATCACACAGTGGCCATCAGCAGTTCACCCACTTTAGACTGTC 397
Db 2006 ICACAGAGAACTTGGAGATCAGGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2085
Qy 398 ATGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 457
Db 2066 CGCATCTGCGAAATCCCTTCCAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2125
Qy 458 AAGAGCTTGGAAATTTTATGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517
Db 2126 AAGACTCAGCAGATGATGATGAGGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2185
Qy 518 AAGAGATGAGGAGTCTATCAGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 577
Db 2186 AAG 2245
Qy 578 CAGCATACCTCAGTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
Db 2246 AAGCATTTTCAATTAAG 2305
Qy 638 CATGACCTGTGCTGCTGAGCTCTCTCTGCTGCTCTTATTAACCTCTCTTATCCGAAA 697
Db 2306 TAGGACGCGGTGATGCTGCTGCTGCTCTCTGCTGCTCTTATTAACCTCTCTTATCCGAAA 2365
Qy 698 TGAAGAGTCT---TCTCTTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 754

Db 2366 TTAAGCGGGCAATGGAGGGAACTGAACAGAGGTACTGTCCATGCTGATGATCCAG 2425
 QY 755 ATGAAGTTCCTTTGGATGAGCACTGTGAGCGGTCCCTTATGATGCGCAGCAAGTGGAGT 814
 Db 2426 ATGAACCTCCCATTTGGAATGAACATTTGAAAGCACTGCTTATGATGCGCAAAATGGAAI 2485
 QY 815 TTGCGGGGAGAGACTTAAACTGGCAATCACTTGGAGAGGGGCTTTTGGAAAGTGG 874
 Db 2486 TCCCCAGAGACCGCTCAAGCTAGTGTAGCTCTTTGGCGTGTGCTTTTGGCCAAATGA 2545
 QY 875 TTCAAGCATCAGCATTTGGCAATTAAGAAATCACTAGTGGCGGAGTGTGGCTGTGGA 934
 Db 2546 TTGAAGCAGATGCTTTTGAATTTGACAGACACACACTTGCAGGACAGTAGCATGCAAA 2605
 QY 935 TCGTGAAGAGGGCGCCAGCGCCAGCAGTACAAATCTGTGATGACTGAGCTTAAATCT 994
 Db 2606 TGTGAAGAGAGGACACACACAGTACAGTACAGCTCTCAATGCTGAATCAAGATCC 2665
 QY 995 TGACCCACATTTGGCCACCACTGAACCTGTAACTGTAACTGTGGAGCTGTGACACAGCAG 1054
 Db 2666 TCATTTCATTTGCTCACCATCTCAATGTGTCAACTTCTAGGTGCTGTACCAAGCCAG 2725
 QY 1055 GAGGCGCTGTGATGTGATTTGTAATCTGCAAAATATGGAATCTTCCAACTACTCA 1114
 Db 2726 GAGGCGCTGTGATGTGATTTGTAATCTTCCAAATTTGGAATCTTCCAACTACTCA 2785
 QY 1115 AGACAAACCTGACTTATTTTCTCAACAGATGCGAGCTACACATGAGCGCTTAAGA 1174
 Db 2786 GGACAGAGAAATGAATTTGCTCCCTACAGCAACCAAGGCGCCAGTCCGTCAAGGGA 2845
 QY 1175 AGAAGAAATGAGCGGCTGGAACAGGCAAGCAACCAAGAGCTAGATAGGCTACCA 1234
 Db 2846 AAGACTACCTTTGGACCAATCCCTG---TGGATCTGAACGCGCTTGGACAGATCTCA 2902
 QY 1235 GCAGGCAAAAGCTTTGGAGCTCGGCTTTACAGAGATGAAATCTGATGTATGAGG 1294
 Db 2903 GTACCCAGAGCTACAGGCTCTGATTTGGAGGAGAGTCTTCACTGATGTATGAGG 2962
 QY 1295 AAGAGAGGATTTCAAGGCTTCAAGGCTTCAAGGAGCCCATCACTATGGAAGATCTGAT 1354
 Db 2963 AAGAGAGGCTTCTGAAGATCTGATATGAGGCTTCTGAGCTTTGGAGATCTCACTG 1302
 QY 1355 ACAGTTTCAAGTGGCCAGAGGATGAGTTCCTGCTTCCAGAAAGTGCATTAAGCGG 1414
 Db 3023 ACAGTCTCAAGTGGCTAAGGCACTGGAGTCTTGGCATCGCGAAGTATCCACAGG 3082
 QY 1415 ACCTGGCAGCGCAAACTTTTATCTGAGACACAGCTGGTGGAGATTTGATATTTG 1474
 Db 3083 ACCTGGCGCAGGAATATCTCTTTTCGGAGAGAACTGTTTAAATCTGTGACTTG 3142
 QY 1475 GCTTGGCCGGATATTTATAAGAACCCCGATTTATGTGAGAAAGAGATATCTGACTTC 1534
 Db 3143 GCTTGGCCGGGATATTTATAAGATCCAGATTTATGTGAGAAAGAGATCTGCGCTTC 3202
 QY 1535 CTCTGAATATGATGCTCCGGAATCTATCTTGACAAATCTACAGCAACAGAGAGAGC 1594
 Db 3203 CTTTGAATATGATGCTCCGGAATCTATCTTGACAAATCTACAGCAACAGAGAGAGC 3262
 QY 1595 TGTGGCTTACAGGATTTGCTGTGGGAATCTTCTCTTCTAGGTCTTCCATACCCAG 1654
 Db 3263 TGTGGCTTACAGGATTTGCTGTGGGAATCTTCTCTTCTAGGTCTTCCATACCCAG 3322
 QY 1655 GAGTACAAATGGAAGAGACTTTTGCAGTCCCTGAGGAGGCAATGAGGATGAGGCTC 1714
 Db 3323 GGGTAAAGATTTGAAGAAATTTGTAGGCGATTTGAAGAGGAGTATGATGAGGCGCC 3382
 QY 1715 CTGAGTACTTACTCTGAAATCTATCAGATCACTGAGCTGTGCGCAGAGAGTCCAA 1774
 Db 3383 CTGATTTACTACACAGAAATGATACAGACCACTGCTGAGCTGTGCGCAGGAGAGCA 3442
 QY 1775 AAGAAAGGCAAGATTTTCAAGAGCTTTGGTAAACTAGGTCTATTTGCTTCAAGCAATG 1834
 Db 3443 GTCAGAGACCCAGTTTTCAGAGTTGGTGAATTTGGAAATCTCTTGAAGATGATG 3502

QY 1935 TACACAGGATGCTTAAGACTACAI 1859
 Db 3503 CTCACAGGATGCGCAAGACTACAI 3527
 RESULT 14
 AX481483
 LOCUS
 DEFINITION Sequence 4071 bp DNA linear PAT 16-AUG-2002
 ACCESSION AX481483
 VERSION AX481483.1 GI:22316397
 FEATURES
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.
 TITLE Method for inhibiting the expression of a target gene
 JOURNAL Patent: WO 02055693-A 97 18-JUL-2002;
 Ribopharma AG (DE)
 FEATURES
 location/Qualifiers
 1. 4071
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1159 a 894 c 1025 g 983 t
 ORIGIN
 Query Match 34.0% Score 713.8; DB 6; Length 4071;
 Best Local Similarity 65.5% Pred. No. 3.8e-197;
 Matches 1077; Conservative 0; Mismatches 562; Indels 6; Gaps 2;
 QY 218 TCAATGATTTTCCCTCAAGATTTCAGGCACCTATGCTGCGAGCCAGGATGTATACA 277
 Db 1886 TTAAGATGATCTTTCAGGACCAAGAGATGATGCTGCTTGCCTCAAGACAGGA 1945
 QY 278 CAGGGGAAGAAATCCICCAAGAAAGAAATTAATCAATCAGATCAGGAAGCACCATACC 337
 Db 1945 CCAAGAAAGACATTTGGTGGTTCAGGCACCTACAGTCTTAGAGCTGTGGCACCACGA 2005
 QY 338 TCTCGGAAACCTCAAGTATCACAGATGCGCATCAGCTGCTTCCACCATTTAGACTGTC 397
 Db 2005 TCACAGGAACCTGGAGATCAGACGACAGATTTGGGNAAGCATCGAAGTCTATGCA 2065
 QY 398 ATGCTAATGTTGCTCCGAGCTTCAGATCACTTGGTTTAAACAAACCAACAAATACAAC 457
 Db 2056 CGGCACTGGGAATCCCTCCACAGATCACTGTTTAAAGATATGAGACCTTGTAG 2125
 QY 458 AAGAGCTTGAATATTTTAGGACCCAGGACGACGCTGTTTATGAAAGATGCACAG 517
 Db 2125 AAGACTCAGGCAATTTTGAAGGATGGAAACCGAAGCTCACTATCCGAGAGTGAGGA 2185
 QY 518 AAGAGGATGAAGTGCTTACACTGCAAGCCACCAACCAAGAGGCTCTGTGGAAGTT 577
 Db 2185 AAGAGGACGAGGAGCTCTACAGCTGCCAGGATGCACTGCTTGGCTGTGCAAGTTG 2245
 QY 578 CAGCATACCTTCATGCTCAAGGACCTTCGACAGTCTTAATCTGGAGCTGATCAGCTTAA 537
 Db 2245 AGGCAATTTTCAATATAGAGGATGCGGAGGAAAGAGAACTTGGAAATCAITTTCTAG 2305
 QY 538 CATGCACTGTGTGGCTGCGACTCTCTTCTGGCTGCTATTTACCTCTTATCCGAAAA 697
 Db 2305 TAGCAGCGGCTGATGCTTCTTCTGGCTTCTTCTTCTGCTTCTTCTTCTTCTTCTTCT 2365
 QY 628 TGAAGAGTCT---TCTTCTGAAATGAAGACTGACTTACTATCAATATTAATGAGCCAG 754
 Db 2365 TTAAGCGGCCCAATGAGGAGGAACTGAAGACAGGCTACTTGTCCATGCTGATGATCCAG 2425
 QY 755 ATGAAGTCTCTTGTGATCAGGAGTGTGAGCGGCTCCCTTATGATCCAGCAAGTGGAGT 814
 Db 2425 ATGACTCCCATTTGATGAACTTTGTAAGAGGAGTCTGCTTATGATGCGCAAGTGGAT 2485

QY 815 TTGCCCCGAGAGACATTAACATGGCAAAATCACTTGGACAGGCGCTTTTGGAAAGTGG 874
 Db 2486 TCCCAAGAGCCGCTGAGCTAGGTAGCTTGGCCGTGGTGTCTTTGGCAAGTGA 2545
 QY 875 TTCAGATACAGATTGGCATTAAGAAATCACTTGGCCGAGCTGTGGCTGTGAAAA 934
 Db 2546 TTGAAGCAGATGGCTTTGGAATTTGACAAGACAGCAACTTGCAGGACAGTACAGCA 2605
 QY 935 TGTCTGAAGAGGGGGCCAGCGGACGAGTACAAAGCTCTGATGACTGAGCTAAATCT 994
 Db 2606 TGTTCAGAGAGAGAGACACACAGCTGAGCACTGAGCTCTGATGACTGAGCTAAATCT 2665
 QY 995 TGACCCCATTTGGCCACCATCTGAACGTGGTTAACTTGGTGGGAACCTGCACAGCAAG 1054
 Db 2666 TCATTCATATTTGGTCAACCACTCAATGGTCAACCTTGTAGTACCTGTACAGCAG 2725
 QY 1055 GAGGCGCTCTGATGGTGTATTTGATGATGATGATGATGATGATGATGATGATGAT 1114
 Db 2726 GAGGCGCACATCTGATGGTGTATTTGATGATGATGATGATGATGATGATGATGATGAT 2785
 QY 1115 AGACCAACAGTGTACTTATTTTCTCAACAGGATGACGACCTTACATGAGTACCAAGA 1174
 Db 2786 GGAGCAGAGAAATGATTTTGTCCCTACAGACCAAGGGGCTTATTCCTCAAGGA 2845
 QY 1175 AGAAAAATGGAGCCAGCCCTGGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1234
 Db 2846 AAGACTAGTGTAGCAATCCCTG---TGGATCTGAAGCGGCTGTGACAGCAAGCA 2902
 QY 1235 GCAGCAGAGCTTGGGCGCTGGCTTTCAGAGAGATGATGATGATGATGATGATGATGAT 1294
 Db 2903 GTAGCCAGCTTGGGCGCTGGCTTTCAGAGAGATGATGATGATGATGATGATGATGAT 2962
 QY 1295 AAGAGAGAGATCTGAGGCTTTCAGAGAGAGATGATGATGATGATGATGATGATGATGAT 1354
 Db 2963 AAGAGAGAGATCTGAGGCTTTCAGAGAGAGATGATGATGATGATGATGATGATGATGAT 3022
 QY 1355 ACAGTTTCAAGTGGCCAGAGCATGAGTGTCTGCTTCCAGAAAGTGCATTTATGCGG 1414
 Db 3023 ACAGTTTCAAGTGGCCAGAGCATGAGTGTCTGCTTCCAGAAAGTGCATTTATGCGG 3082
 QY 1415 ACCTGGCAGCAGAAACATCTTATCTAGAGAGAGATGATGATGATGATGATGATGATGAT 1474
 Db 3083 ACCTGGCAGCAGAAACATCTTATCTAGAGAGAGATGATGATGATGATGATGATGATGAT 3142
 QY 1475 GCCTTGGCCGGATATTTAAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 1534
 Db 3143 GCCTTGGCCGGATATTTAAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 3202
 QY 1535 CTCTGAATGAGTGGCTCCCAATCTATCTTACAGAAATCTAGAGAGAGATGATGATGATGAT 1594
 Db 3203 CTCTGAATGAGTGGCTCCCAATCTATCTTACAGAAATCTAGAGAGAGATGATGATGATGAT 3262
 QY 1595 TGTGGCTTACGAGATTTGCTGGGAAATCTCTCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 1654
 Db 3263 TGTGGCTTACGAGATTTGCTGGGAAATCTCTCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 3322
 QY 1655 GAGTACAAATGATGAGGATTTTGGAGTGGCTGAGAGAGAGATGATGATGATGATGATGAT 1714
 Db 3323 GAGTACAAATGATGAGGATTTTGGAGTGGCTGAGAGAGAGATGATGATGATGATGATGAT 3382
 QY 1715 CTGAGTACTTACTCTCTGAAATCTATCAGATCACTGAGGATGATGATGATGATGATGATGAT 1774
 Db 3383 CTGAGTACTTACTCTCTGAAATCTATCAGATCACTGAGGATGATGATGATGATGATGATGAT 3442
 QY 1775 AAGAAAGGCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTG 1834
 Db 3443 GTCAGAGAGAGAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGATTTGCAAGAT 3502
 QY 1835 TACAGAGAGATGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1899
 Db 3503 CTGAGAGAGATGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3572

RESULTS

AF063658

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

FEATURES

SOURCE

AF063658 Homo sapiens vascular endothelial growth factor receptor 2 (KDR)

AF063658 mRNA, complete cds.

AF063658 AF063658.1 GI:3132832

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4071)

Coding region for human VEGF receptor KDR (VEGFR-2)

Unpublished

2 (bases 1 to 4071)

Y.Y., Whitney, R.G. and Sato, J.D.

Direct Submission

Submitted (05-MAY-1998) Adirondack Biomedical Research Institute.

10 Old Barn Rd., Lake Placid, NY 12946, USA

Location/Qualifiers

1. 4071

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="endothelium"

/issue_type="umbilical vein"

1. 4071

/gene="KDR"

1. 4071

/note="KDR"

/note="VEGFR-2; tyrosine kinase"

/codon_start=1

/product="vascular endothelial growth factor receptor 2"

/protein_id="AAC16450.1"

/db_xref="GI:3132832"

/translation="MESKVLIAVALWCVETRAASVCLPVSULDPLRISIKDILITK

ANTLIQTCGRDLNEMFNQSSQEVETECSDGLFCKTLTIPIKVIDNIGAYK

CYRIDLASVIVVYDYSFPIASVDHGVYITENKTVIPCLGISNLNYS

LCARPERFVPCNRIKSWDKGFTIPSYMISYAGVFCARINDESVQIMIVV

VGYRIYDVLSPSHGIELSVGKLVNLTARIELNVGIDFNWEPSSKHOKLVNED

LITQSGEMKKFLSTLITUGVTRSDQGLNCAASSGLMTKNSITFVRHEKPFVAGS

GMESIVATVGERVIRPAKVICYPPEIKYKNGIPLESNHTIKAGHVLIIMEVSRD

TGNIVILNPISEKOSHVSLVYVPPQIGESLISPVDSYQYGTITLITCIVAI

PPPHIHWOLEECANEFSAQSVNTPYPCESRSVEDFGGKNIEVKNQFALIE

GKNITVTLIQANYSALYKCAVNVKGERVIFSHVTRGPEITLQPDQKTEQES

VSLMCTADRSTFENLWYKLPQPLIHVGLPVPCKNLDLWLNAMFNSINDI

LIMELKNASIQDQGVCLAODRKYKRCVVRQLTVLERVAPTIIGLNQVTSIGE

SIEVCTASGNPPPMWFKDNETLVEDSGTVLKDGNRLTIRVRKEDSLTCQPC

SVLSCAVEAFFIIEGAQEKINLEIILVGTAVIAMFFWLLVILIRIVKANGELK

TGYLSIVMDELPLDEHCELRPYDASKWEPDRRLKGLKPLGRAGQGVTEADFGI

DKATIRIVAVKMKEGATHSEHRLMSKLIHGHHLNVNLLGACTKPGGLMV

IVETCFGNLSYLRKNEFYKTKGARGKQGVGAIPDLKRLDLSITSSQSS

ASSGFVEEKLSVDEEAPDLKYDLTLEHLICYSFOVAKGMEELASRKLIRDLA

ARNILLSEKNVWICDFGLARDIYKDPDYVKGDARLPKWMAPETIDRVYLIQSDY

WSFVLLWEIFSLGASPYPCVADIEECRLKGTGEMRAPDTTETMTOLMCWGE

BSQPTSELVHLNLLQANQQDQDIIVLPISLSEEDSLSLPTSPSCWGE

EVGCPHFHNTAGISQTLQNSRKRPVSKTFIDIEPEEYKVPIDNDQDGM

VLASEELILREDRIKLSPSGFMVPSKRSVASGNSQITSGYSGVHSDIDITVVS

SEAEKLKLEIGVGTGTATQIAPDSGTTLSPEPV"

BASE COUNT 1169 a 894 c 1025 g 983 t

ORIGIN

Query Match

Score 713.8; DB 9; Length 4071;

Best Local Similarity 55.5%; Pred. No. 3.8e-137;

Matches 1077; Conservative 0; Mismatches 562; Indels 6; Gaps 2;

Cj 218 TCAATGATGCTTCCTCCAGAGAGAAATTAACATCAGAGATCAGAGACCATACC 337

Db 1886 TTAGATGCTTCCTCCAGAGAGAAATTAACATCAGAGATCAGAGACCATACC 337

Qy 278 CAGGGAAGAAATTCCTCCAGAGAGAAATTAACATCAGAGATCAGAGACCATACC 337

Db 1946 CCAAGAAAAGACATTCGGTGGTGGAGCAGCTCACAGTCTAGAGCGGTGGCCACCACGA 2025
QY 338 TCCTGCGAAACCTCAGTCATGACACACAGTGGCCATCAGCAGTTCACACACTTTAGACTGTC 397
Db 2006 TCACAGGAACCTCGAGAAATCAGACGACAGTATTTGGGAAAGCATCGAAGTCTCATGCA 2055
QY 398 ATGCTAATGGTGTGCGGAGGCTCAGATCAGTGGTTTAAACCAACCAAAATACAC 457
Db 2066 CGGCATCTGGGAATCCCTCCACAGATCATGGTTTAAAGATATGAGACCTTTAG 2125
QY 458 AAGAGCTGGATTAATTTAGGACAGGAGCAGCAGCTTTTATGAAAGATCAGAC 517
Db 2126 AAGACTCAGGATGTATGAGGATGGGACGGAACCTCAGTATCGCAGAGTGGA 2185
QY 518 AAGACATGAAGTGTCTATCAGTCAAAATCCACCAACCAAGGCTCTGTGGAAAGT 577
Db 2186 AAGAGACAGAGGCTCTACACCTTGCAGGATGATGAGTGTCTTGGCTGTGAAAGTGG 2245
QY 578 CAGCATACCTCTACTGTTCAAGGAACCTCGGACAGAGTCTAATCTGGAGCTGATCAGCTTAA 637
Db 2246 AGGCATTTTCATATAGAAAGTGGCCAGGAAAGACAGCACTTGGAAATCATTTCTAG 2305
QY 638 CATGCACCTGTGGTGGCAGCTCTCTTGGCTCTTAAACCTCTTATCCGAAAA 697
Db 2306 TAGGCAGGCGGTGATGGCCATGCTCTGCTACTCTGTGATCATCTCTACGACCG 2365
QY 698 TGAAGAGTCTCTTCTTGTGAATTAAGATGAGTCTACCTTCAATTAATATGGACCG 754
Db 2366 TTAAGCGGCAATGGAGGGGACATCAAGCAGGCTACTTGTCCATCTGTATGGAATCC 2425
QY 755 ATGAAGTCTCTTGGATGAGCAGTGTGAGTGGCTCCCTTATGATGACCAAGTGGAGT 814
Db 2426 ATGAAGTCTCTTGGATGAGCAGTGTGAGTGGCTCCCTTATGATGACCAAGTGGAGT 874
QY 815 TTGCGCGGAGAGACTTAAATGGGCAATCAGTTCGAGAGGGGCTTTTGAAGATGG 874
Db 2486 TCCCGAGAGCGGTGAGCTAGTATGAGTCTTGGCGGTGGTGGCTTTGGCCAAAGTGA 2545
QY 875 TTCAGCTATGAGCTTTGGCAATTAAGAAATCACTGCTGCGGAGTGTGGCTGTGAAA 934
Db 2546 TTGAAGAGAGTGGCTTTGGCAATTAAGAAATCACTGCTGCGGAGTGTGGCTGTGAAA 2605
QY 935 TGCTGAAGAGGGGCGCCAGGCGGAGTACAGAGTCTGATGAGTGAAGTGAAGTGAAGT 994
Db 2606 TGTGGAAGAGGAGCAACACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2665
QY 995 TGACCCATTTGCCACCACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 3054
Db 2666 TCATTCATATTTGTCACCACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 2725
QY 1055 GAGGCGCTCTGATGGTGTGATGTTGATCTGCAATTAAGAAATCTTCCAACTACCTCA 1114
Db 2726 GAGGCGCACTCTGATGGTGTGATGTTGATCTGCAATTAAGAAATCTTCCAACTACCTCA 2785
QY 1115 AGAGCAATGCGAGTATTTTCTCAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1174
Db 2786 GGAGCAGAGAAATGAATTTGTCCTTACAGGACCAAGAAAGGAGGAGGAGGAGGAGGAGGAG 2845
QY 1175 AAGAAAAATGAGAGCGGCTTGAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1234
Db 2846 AAGACTACGTTGAGCAATCTCTG---TGGATCTGAAGCGGCTTGGAGAGTATGATAGAG 2902
QY 1235 CGAGCGAAGCTTTGGAGCTTCGGGCTTTTCAAGAGATAAAAGTCTGAGTATGTTGAGG 1294
Db 2903 GTACCCAGAGCTCAGCAGCTCTGGATTTTGGAGGAGAGTCTCTGAGTATGATAGAG 2962
QY 1295 AAGAGGAGGATCTGACGCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1354
Db 2963 AAGAGGAGGATCTGAGATCTGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3022
QY 1355 ACAGTTTCAAGTGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1414
Db 3023 ACAGTTTCAAGTGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3082

QY 1415 ACCGCGCAGGAGAAACATCTTTTATCTGAGAAACAGCTGCTGAGAGTTTGTGATTTG 1474
Db 3083 ACCGCGCAGGAGAAACATCTTTTATCTGAGAAACAGCTGCTGAGAGTTTGTGATTTG 3142
QY 1475 GCTTTCGCGGAGATTTTATAGAAACCGGATTTGAGAAAGAGATGATCTGACATTC 1534
Db 3143 GCTTTCGCGGAGATTTTATAGAAACCGGATTTGAGAAAGAGATGATCTGACATTC 3202
QY 1535 CTCTGAATGGATGCTGCTGCGGATTTTATCTGAGAAATCTACAGCAACAGAGCGAG 1594
Db 3203 CTCTGAATGGATGCTGCTGCGGATTTTATCTGAGAAATCTACAGCAACAGAGCGAG 3262
QY 1595 TGTGCTTTAGGAGATGCTGCTGCGGATTTTATCTGAGAAATCTACAGCAACAGAGCGAG 1654
Db 3263 TGTGCTTTAGGAGATGCTGCTGCGGATTTTATCTGAGAAATCTACAGCAACAGAGCGAG 3322
QY 1655 GATGCAATGAGTGGAGTCTGCTGCGGATTTTATCTGAGAAATCTACAGCAACAGAGCGAG 1714
Db 3323 GATGCAATGAGTGGAGTCTGCTGCGGATTTTATCTGAGAAATCTACAGCAACAGAGCGAG 3382
QY 1715 CTGAGTCTCTACTCTGCTGAGAAATCTACAGCAACAGAGCGAGTCTGCTGCGGAGCGAG 1774
Db 3383 CTGAGTCTCTACTCTGCTGAGAAATCTACAGCAACAGAGCGAGTCTGCTGCGGAGCGAG 3442
QY 1775 AAGAAAGCGGAGATTTGAGAAATCTGAGAAATCTGAGAAATCTGAGAAATCTGAGAAAT 1834
Db 3443 AAGAAAGCGGAGATTTGAGAAATCTGAGAAATCTGAGAAATCTGAGAAATCTGAGAAAT 3502
QY 1835 TSCAACAGGATGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1899
Db 3503 TSCAACAGGATGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3527

Search completed: December 8, 2002, 11:59:20
Job time: 6945.36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 04:25:08 ; Search time 111.05 Seconds
(without alignments)
2048.198 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600

Perfect score: 101

Sequence: 1 ataggctgtgtgtctcc.....tatgtaaatgtttcaagtt 101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

tal number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_101002.*

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	101	100.0	1256	22	AAD19456	Human flt-1 gene D
2	32	31.7	288	23	AAS51865	Staphylococcus aur
3	32	31.7	300	19	AAV45103	Staphylococcus aur
4	32	31.7	300	19	AAV45098	Staphylococcus aur
5	32	31.7	300	19	AAV33271	Staphylococcus aur
6	32	31.7	300	20	AAV56688	S. aureus RAIC cod
7	32	31.7	303	23	AAS45756	Staphylococcus aur
8	32	31.7	303	23	AAS55130	Staphylococcus aur
9	32	31.7	303	23	AAS55306	Staphylococcus aur

C 10	32	31.7	3400	21	AA251225	Staphylococcus aur
C 11	32	31.7	13542	18	AAV74465	Staphylococcus aur
C 12	31.6	31.3	160771	24	ABQ88179	Human osteoblast d
C 13	31.4	31.1	36778	23	ABL07548	Drosophila melanog
C 14	30.2	29.9	4328	23	ABL27278	Drosophila melanog
C 15	30.2	29.9	7514	23	ABL04220	Drosophila melanog
C 16	30	29.7	5863	23	ABL04354	Drosophila melanog
C 17	29.8	29.5	17979	22	AAK64964	Human immune/haema
C 18	29.8	29.5	17979	22	AAK71684	Human immune/haema
C 19	29.4	29.1	1462	22	AAI81101	Human polynucleoti
C 20	29.4	29.1	580	22	ABK36027	cDNA sequence #418
C 21	29.4	29.1	788	22	AAI95236	Human neuroblastom
C 22	29.4	29.1	788	22	AAI98133	Human neuroblastom
C 23	29.4	29.1	3707	22	AAO08367	Human secreted pro
C 24	29.4	29.1	429	24	ABT03068	Human breast speci
C 25	29	28.9	5598	24	ABL33635	Human immune syste
C 26	28.6	28.5	2227	24	ABK14577	Human cDNA encodin
C 27	28.8	28.5	6379	22	AAI46347	Tumour suppressor
C 28	28.6	28.3	415	22	AAF67314	Novel human polynu
C 29	28.6	28.3	6686	24	ABL32460	Human immune syste
C 30	28.6	28.3	75855	24	ABK65261	Human genomic DNA
C 31	28.6	28.3	147724	24	ABK83566	Human cDNA differe
C 32	28.4	28.1	62	22	AAI96700	Human neuroblastom
C 33	28.2	27.9	379	24	ABL61747	Colon adenocarcino
C 34	28.2	27.9	498	23	ABV57525	Human prostate exp
C 35	28.2	27.9	1212	21	AAZ65332	Human secreted pro
C 36	28.2	27.9	1608	20	AAZ00815	Human secreted pro
C 37	28.2	27.9	10020	24	ABL34292	Human immune syste
C 38	28.2	27.9	12356	22	AAI46510	Tumour suppressor
C 39	28	27.7	195	21	AAI17824	Human secreted pro
C 40	28	27.7	380	22	ABA57410	Human foetal liver
C 41	28	27.7	380	22	AAK05446	Human bone expres
C 42	28	27.7	380	22	AAK31045	Human bone marrow
C 43	28	27.7	380	22	AAI36955	Probe #5611 used t
C 44	28	27.7	380	24	ABS05795	Human genome-deriv
C 45	28	27.7	5000	22	AAI4506	Human GSTT1 deleti

ALIGNMENTS

RESULT 1
AAD19456
ID: AAD19456 Standard: DNA; 1256 BP.
XX
XX
AC AAD19456;
DI
DI 18-DEC-2001 (First entry)
XX
XX
DE Human flt-1 gene DNA fragment #5.

XX Human; flt-1 gene; cytostatic; haemostatic; gene mapping; gynaecological;
XX Ophthalmological; rheumatoid arthritis; endometriosis; arthrogenic disease;
XX diabetic retinopathy; psoriasis; VEGFR-1; drug therapy; pharmacogenetic;
XX vascular endothelial growth factor; single nucleotide polymorphism; SNP;
XX Cancer; medication; ds.
OS Homo sapiens.

Key	Location/Qualifiers
FI intron	1:579
FI exon	/*tag= a
FI intron	580..664
FI intron	/*tag= b
FI intron	665..1256
FI variation	/*tag= c
FI variation	replace (696, C)
FI variation	/*tag= d
FI variation	/*standard_name= "Single nucleotide polymorphism (SNP)"

EP1130123-A2.
XX
XX 05-SEP-2001.

CC conjunctivitis, osteomyelitis, skin infections and toxic shock syndrome.
 CC The present sequence is also useful for detecting Staphylococcal
 CC infections in biological samples.
 XX
 SQ Sequence 3400 BP; 1256 A; 510 C; 648 G; 986 T; 0 other;
 Query Match 31.7%; Score 32; DB 21; Length 3400;
 Best Local Similarity 68.8%; Pred. No. 5.1; Indels 0; Gaps 0;
 Matches 44; Conservative 0; Mismatches 20; Labels 0; Gaps 0;
 QY 38 ATGCTTTCCTGCTGATATCTTCTCTTCATTAAGTAAGTAAAGCTTCA 97
 IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 Db 209 ATGCTTTCATGTTGGCCATCTTCCTGCTTCTCAGAGAAATGAGAGCTTCA 150
 QY 98 AGTT 101
 II I
 Db 149 AGAT 146
 RESULT 11
 /74465
 AAV74465 standard; DNA; 13542 BP.
 XX AAV74465;
 AC
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE Staphylococcus aureus contig SEQ ID #154.
 XX
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 XX
 OS Staphylococcus aureus.
 XX
 FH Location/Qualifiers
 FI misc_feature 781..840
 FI /*tag= a
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 FI 2581..2640
 FI /*tag= b
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 FI 4381..4440
 FI misc_feature 4381..4440
 FI /*tag= c
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 FI 6181..6240
 FI misc_feature 6181..6240
 FI /*tag= d
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 FI 7981..8040
 FI misc_feature 7981..8040
 FI /*tag= e
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 FI 9781..9840
 FI misc_feature 9781..9840
 FI /*tag= f
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 FI 160771 BP.

FI misc_feature 11581..11640
 FI /tag= g
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 FI 13381..13440
 FI misc_feature 13381..13440
 FI /tag= h
 FI /note= "these bases represent a line of missing text in
 FI the sequence listing in the specification. They
 FI are included to maintain the nucleotide numbering
 FI given in the specification for this DNA sequence."
 XX
 PX EP786519-A2.
 XX
 PD 30-JUL-1997.
 XX
 XX
 PF 07-JAN-1997; 57EP-0100117.
 XX
 PX 05-JAN-1996; 56US-0609861.
 XX
 PX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 PI
 DR WPI; 1997-374522/35.
 XX
 XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 PI stored on computer readable medium and used in the production of
 PI anti-S.aureus vaccines
 XX
 PS Claim 1; Page 776-784; 327pp; English.
 XX
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock,
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 XX
 SQ Sequence 13542 BP; 3969 A; 2351 C; 1925 G; 4813 T; 484 other;
 Query Match 31.7%; Score 32; DB 18; Length 13542;
 Best Local Similarity 68.8%; Pred. No. 6.8;
 Matches 44; Conservative 0; Mismatches 20; Labels 0; Gaps 0;
 QY 38 ATGCTTTCCTGCTGATATCTTCTCTTCATTAAGTAAGTAAAGCTTCA 97
 IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 Db 9494 ATGCTTTCATGTTGGCCATCTTCCTGCTTCTCAGAGAAATGAGAGCTTCA 9553
 QY 98 AGTT 101
 II I
 Db 9554 AGAT 9557
 RESULT 12
 ASQ88179/c
 ID ASQ88179 standard; CDNA; 160771 BP.
 XX
 AC ASQ88179;

XX WO200171042-A2.
XX
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PI interactions -
XX
XX Claim 1; SEQ ID NO 33307; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (AB57737-AB572072).
CC (AB57737-AB572072).
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4328 BP; 1144 A; 996 C; 1077 G; 1111 T; 0 other;
SQ
Query Match 29.9%; Score 30.2; DB 23; Length 4328;
Best Local Similarity 62.7%; Pred. No. 18;
Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 20 CTGACTCATTAGAGTGGATGCTTTTCCTGCTGTGATATTCCTGCTTCATTA 79
DB 3550 CTGGTAAAGAGTGTAGCCCTTACTTTTGGAAAAGTCTGTTCTACATTT 3609
QY 80 GATATGTAATGCTTT 94
DB 3610 ATAATATTAATTTT 3624
SULT 15
ABL04220
ID ABL04220 standard; cDNA; 7914 BP.
XX
XX ABL04220;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 7142.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX

PA (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR P-PSDB; ABB60117.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PI interactions -
XX
XX Claim 1; SEQ ID NO 7142; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (AB57737-AB572072).
CC (AB57737-AB572072).
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 7514 BP; 1679 A; 2004 C; 1981 G; 2250 I; 0 other;
SQ
Query Match 29.9%; Score 30.2; DB 23; Length 7514;
Best Local Similarity 62.7%; Pred. No. 21;
Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 20 CTGACTCATTAGAGTGGATGCTTTTCCTGCTGTGATATTCCTGCTTCATTA 79
DB 906 CTGGTAAAGAGTGTAGCCCTTACTTTTGGAAAAGTCTGTTCTACATTT 965
QY 80 GATATGTAATGCTTT 94
DB 966 ATAATATTAATTTT 980
Search completed: December 8, 2002, 06:02:36
Job time : 134.05 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 05:16:46 ; Search time 801.97 Seconds
(without alignments)
3665.203 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600

Perfect score: 101

Sequence: 1 atattgtccgtgtgtctcc.....tatgtaaatgcttccaagtt 101

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

total number of hits satisfying chosen parameters: 4109290

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_un.*
- 28: em_vi.*
- 29: em_vt.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rdt.*
- 36: em_htg_mam.*
- 37: em_htg_vit.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	99.4	98.4	164519	9	AL138712 Human DNA
C 2	99.4	98.4	164519	2	AC016311 Homo sapi
C 3	97.2	56.6	101587	2	AC128808 Rattus no
C 4	96.6	36.2	170443	2	AC116776 Mus muscu
C 5	96.2	35.8	96559	9	AC113610 Homo sapi
C 6	95.4	35.0	172076	2	AC121895 Mus muscu
C 7	95.4	35.0	25813	2	AC094998 Rattus no
C 8	95.2	34.9	118110	9	AC098158 Homo sapi
C 9	95.2	34.9	15223	2	AC124802 Homo sapi
C 10	95.2	34.9	15223	2	AC124802 Homo sapi
C 11	95.2	34.9	15223	2	AC124802 Homo sapi
C 12	95.2	34.9	15223	2	AC124802 Homo sapi
C 13	95.2	34.9	15223	2	AC124802 Homo sapi
C 14	95.2	34.9	15223	2	AC124802 Homo sapi
C 15	95.2	34.9	15223	2	AC124802 Homo sapi
C 16	95.2	34.9	15223	2	AC124802 Homo sapi
C 17	95.2	34.9	15223	2	AC124802 Homo sapi
C 18	95.2	34.9	15223	2	AC124802 Homo sapi
C 19	95.2	34.9	15223	2	AC124802 Homo sapi
C 20	95.2	34.9	15223	2	AC124802 Homo sapi
C 21	95.2	34.9	15223	2	AC124802 Homo sapi
C 22	95.2	34.9	15223	2	AC124802 Homo sapi
C 23	95.2	34.9	15223	2	AC124802 Homo sapi
C 24	95.2	34.9	15223	2	AC124802 Homo sapi
C 25	95.2	34.9	15223	2	AC124802 Homo sapi
C 26	95.2	34.9	15223	2	AC124802 Homo sapi
C 27	95.2	34.9	15223	2	AC124802 Homo sapi
C 28	95.2	34.9	15223	2	AC124802 Homo sapi
C 29	95.2	34.9	15223	2	AC124802 Homo sapi
C 30	95.2	34.9	15223	2	AC124802 Homo sapi
C 31	95.2	34.9	15223	2	AC124802 Homo sapi
C 32	95.2	34.9	15223	2	AC124802 Homo sapi
C 33	95.2	34.9	15223	2	AC124802 Homo sapi
C 34	95.2	34.9	15223	2	AC124802 Homo sapi
C 35	95.2	34.9	15223	2	AC124802 Homo sapi
C 36	95.2	34.9	15223	2	AC124802 Homo sapi
C 37	95.2	34.9	15223	2	AC124802 Homo sapi
C 38	95.2	34.9	15223	2	AC124802 Homo sapi
C 39	95.2	34.9	15223	2	AC124802 Homo sapi
C 40	95.2	34.9	15223	2	AC124802 Homo sapi
C 41	95.2	34.9	15223	2	AC124802 Homo sapi
C 42	95.2	34.9	15223	2	AC124802 Homo sapi
C 43	95.2	34.9	15223	2	AC124802 Homo sapi
C 44	95.2	34.9	15223	2	AC124802 Homo sapi
C 45	95.2	34.9	15223	2	AC124802 Homo sapi

ALIGNMENTS

RESULT 1
AL138712/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-502P18 on chromosome 13.
complete sequence.
ACCESSION AL138712
VERSION AL138712.15 GI:15131448
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 164519)
AUTHORS Bates.K.
TITLE Direct Submission

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinn, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, J., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, N., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgeson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homs, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, H.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louis, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, J., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nockenkwo, S., Oguh, M., Okunade, G.,
 Oquaye, N., Oriado, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Re, I.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saver, S.,
 Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
 Sodergren, E., Sosa, T., Sparks, A., Stanley, H., Store, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tareisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalobos, D., Vinson, R., Wang, S.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wiczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 101987)
 Worley, K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center -----
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information -----
 Center project name: KAMN
 Center clone name: CH230-69422
 ----- Summary Statistics -----
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 55655 bases at least Q40
 Consensus quality: 59829 bases at least Q30
 Consensus quality: 62415 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length.
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 59 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1002: contig of 1002 bp in length
 1003 1102: gap of unknown length
 1103 2129: contig of 1027 bp in length
 2130 2229: gap of unknown length
 2230 3237: contig of 1008 bp in length
 3238 3337: gap of unknown length
 3338 4784: contig of 1447 bp in length
 4785 4884: gap of unknown length
 4885 6614: contig of 1730 bp in length
 6615 6715: gap of unknown length
 6716 8257: contig of 1543 bp in length
 8258 9401: contig of unknown length
 9402 9501: contig of 1044 bp in length
 9502 10661: contig of 1160 bp in length
 10662 10761: gap of unknown length
 10762 11930: contig of 1169 bp in length
 11931 12030: gap of unknown length
 12031 13552: contig of 1522 bp in length
 13553 13632: gap of unknown length
 13633 15262: contig of 1610 bp in length
 15263 15363: gap of unknown length
 15364 16603: contig of 1141 bp in length
 16604 17564: contig of unknown length
 17565 18044: gap of unknown length
 18045 19537: contig of 1473 bp in length
 19538 19637: gap of unknown length
 19638 21165: contig of 1528 bp in length
 21166 21265: gap of unknown length
 21266 22824: contig of 1959 bp in length
 22825 23924: gap of unknown length
 23925 23965: contig of 1041 bp in length
 23966 24065: gap of unknown length
 24066 25710: contig of 1645 bp in length
 25711 25811: gap of unknown length
 25812 27247: gap of unknown length
 27248 28660: contig of 1613 bp in length
 28661 28963: contig of 1003 bp in length
 28964 30063: gap of unknown length
 30064 31123: contig of 1060 bp in length
 31124 31223: gap of unknown length
 31224 32667: contig of 1644 bp in length
 32668 32967: gap of unknown length
 32968 34070: contig of 1103 bp in length
 34071 34170: gap of unknown length
 34171 36370: contig of 2200 bp in length
 36371 37602: contig of 1331 bp in length
 37603 37901: gap of unknown length
 37902 39087: contig of 1186 bp in length
 39088 39187: gap of unknown length
 39188 40712: contig of 1525 bp in length
 40713 42242: gap of unknown length
 42243 42342: gap of unknown length
 42343 43567: contig of 1225 bp in length
 43568 43667: gap of unknown length
 43668 44985: contig of 1318 bp in length
 44986 45085: gap of unknown length
 45086 46200: contig of 1115 bp in length
 46201 46300: gap of unknown length
 46301 47472: contig of 1172 bp in length
 47473 47572: gap of unknown length
 47573 49428: contig of 1856 bp in length
 49429 49528: gap of unknown length
 49529 51428: contig of 1900 bp in length
 51429 51528: gap of unknown length
 51529 52659: contig of 1171 bp in length
 52660 52760: gap of unknown length
 52761 54476: contig of 1677 bp in length
 54477 54576: gap of unknown length
 54577 55912: contig of 1336 bp in length
 55913 56012: gap of unknown length
 56013 57709: contig of 1697 bp in length
 57710 57809: gap of unknown length
 57811 59785: contig of 1980 bp in length

```

* 1 953: contig of 953 bp in length.
* 954 1053: gap of 100 bp
* 1054 7617: contig of 8564 bp in length.
* 7618 16955: gap of 100 bp
* 16956 17055: contig of 9238 bp in length
* 17056 28121: gap of 100 bp
* 28122 28221: contig of 11066 bp in length
* 28222 60294: gap of 100 bp
* 60295 60394: contig of 32073 bp in length
* 60395 95569: gap of 100 bp
* 95570 95669: contig of 35175 bp in length
* 131655 131754: gap of 100 bp
* 131755 170443: contig of 35985 bp in length
* 170444 32488: gap of 100 bp
* 32489 32968: contig of 3689 bp in length.

```

FEATURES

```

source
1. 170443
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-47245"
/clone.lib="RP24-47245 Male Mouse BAC"

misc_feature
1. 953
/note="assembly_fragment"

misc_feature
1054. 7617
/note="assembly_fragment"

misc_feature
7718. 16955
/note="assembly_fragment"

misc_feature
17056. 28121
/note="assembly_fragment"

misc_feature
28222. 60294
/note="assembly_fragment"

misc_feature
60395. 95569
/note="assembly_fragment"

misc_feature
95670. 131654
/note="assembly_fragment"

misc_feature
131755. 170443
/note="assembly_fragment"

BASE COUNT 54653 a 32488 c 32968 g 49630 t 704 others
ORIGIN

```

```

Query Match 36.28; Score 36.6; DB 2; Length 170443;
Best Local Similarity 62.68; Pred. No. 3.5;
Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QV 10 CTGTGTCACCATTCATGATTAACCTGAGGCTTACCTGCTGATTAATTCCTGCT 69
||||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62057 CTGTGTCACCATTCATGATTAACCTGAGGCTTACCTGCTGATTAATTCCTGCT 61998
||||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 70 TTCTCATGATGATGCTAAAGCTTTCAAGT 100
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61997 TTCTCATGATGATGCTAAAGCTTTCAAGT 61967
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 5
AC113610/C
LOCUS AC113610 Homo sapiens BAC clone RP11-373H2 from 2, complete sequence.
DEFINITION AC113610
ACCESSION AC113610.4 GI:21694035
VERSION AC113610.4
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 96559)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE
2 (bases 1 to 96559)
AUTHORS Abbott, S., Haakenson, W. and Creason, K.
TITLE The sequence of Homo sapiens BAC clone RP11-373H2
JOURNAL Unpublished (2001)

```

```

REFERENCE
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (04-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center Code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0373H02

```

NOTES: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30). An attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatem, C.M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and co-workers at <http://www.chori.org> VEC13R: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-62516, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-373H2. Actual end is at base position 63477 of RP11-62516.

The region from 20416 to 20419 is covered only by a per product from clone DNA.

```

FEATURES
Source
1. 96559
/organism="Homo sapiens"
/db_xref="taxon:9606"

```


Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K.,
 Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C.,
 Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, S.,
 Hagos, B., Heard, A., Horton, L., Howland, J. C., Jones, C., Kana, I.,
 Karatas, A., Lehocsky, J., Lieu, C., Locke, K., Macdonald, P.,
 Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
 Melidrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
 Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
 Peterson, K., Pollara, V., Rile, R., Roberts, D., Roy, A., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
 Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
 Wheeler, J., Wu, X., Wyman, D., Ye, W. J., and Zody, M.

Direct Submission

Submitted (28-JUL-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 148110)

REFERENCE

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginder, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Karatas, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
 Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Melidrim, J., Meneus, I.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, A.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talmes, J., Tesfaye, S., Theodor, J.,
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

Direct Submission

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 148110)

REFERENCE

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginder, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Karatas, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
 Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Melidrim, J., Meneus, I.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, A.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talmes, J., Tesfaye, S., Theodor, J.,
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

Direct Submission

Submitted (24-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 24, 2002 this sequence version replaced g118610683.

All repeats were identified using RepeatMasker:

Smit, A. F. A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

TITLE

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research.
 Center code: WTHR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L552
 Center clone name: 42_F_20

FEATURES

Source

Source	Location/Qualifiers
1. 148110	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="17"
	/map="17"
	/clone="RP11-42F20"
	/clone_lib="RP11 Human Male BAC"
repeat_region	344..1086
	/rpt_family="Charliela"
repeat_region	complement(1201..1280)
	/rpt_family="MER2"
repeat_region	1281..1389
	/rpt_family="AluX"
repeat_region	complement(1590..1717)
	/rpt_family="MER2"
repeat_region	1721..2327
	/rpt_family="Charliela"
repeat_region	complement(2593..2901)
	/rpt_family="AluY"
repeat_region	complement(3238..3536)
	/rpt_family="AluSc"
repeat_region	complement(3546..4392)
	/rpt_family="SVA"
repeat_region	4461..4632
	/rpt_family="L1MC4a"
repeat_region	4745..5033
	/rpt_family="AluX"
repeat_region	5034..5053
	/rpt_family="(TA)n"
repeat_region	5076..5193
	/rpt_family="L2"
repeat_region	5233..5328
	/rpt_family="L1MC4a"
repeat_region	complement(5329..5465)
	/rpt_family="FLAM_C"
repeat_region	5466..5560
	/rpt_family="L1MC4a"
repeat_region	5585..5868
	/rpt_family="AluJo"
repeat_region	complement(6178..6555)
	/rpt_family="L1R33A"
repeat_region	complement(6881..7172)
	/rpt_family="AluX"
repeat_region	complement(7392..7941)
	/rpt_family="L1PA13"
repeat_region	complement(7982..8242)
	/rpt_family="MER38B"
repeat_region	8243..8533
	/rpt_family="AluX"
repeat_region	complement(8534..8608)
	/rpt_family="MER38B"
repeat_region	9360..11422
	/rpt_family="L1PA13"
repeat_region	11475..11721
	/rpt_family="AluY"
repeat_region	11732..12006
	/rpt_family="AluJo"
repeat_region	13948..14228
	/rpt_family="L1MC4a"
repeat_region	14590..14625
	/rpt_family="AT-rich"
repeat_region	14642..14673
	/rpt_family="(T)n"
repeat_region	15048..15355

LD 73 TTGGCGAACGGGA 60
|||||

RESULT 7

CNS04EJU/c

LOCUS

DEFINITION

Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone

104B12 of library G from Tetraodon nigroviridis, genomic survey

sequence.

AL287139.1

VERSION

AL287139.1 GI:8025621

KEYWORDS

GSS; genome survey sequence.

SOURCE

Tetraodon nigroviridis.

ORGANISM

Tetraodon nigroviridis.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

Tetraodontidae; Tetraodon.

1 (bases 1 to 888)

Roest-Crollius H., Jallion O., Dasilva C., Bouneau L., Fisher C.,

Bernot A., Pizares C., Wincker P., Brottier P., Quetier F.,

Saurin W. and Weissenbach J.

Human gene number estimate provided by genome wide analysis using

Tetraodon nigroviridis DNA sequence

Unpublished

2 (bases 1 to 888)

Roest-Crollius H., Jallion O., Dasilva C., Pizares C., Fisher C.,

Bouneau L., Billault A., Quetier F., Saurin W., Bernot A. and

Weissenbach J.

Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

Unpublished

3 (bases 1 to 888)

Genoscope.

Direct Submission

Submitted (12-APR-2000)

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

http://www.genoscope.cns.fr/tetraodon.

FEATURES

Location/Qualifiers

1..888

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone_lib="104B12"

/clone_lib="G"

/note="Genoscope sequence ID : CGB310-DA06SP1-end"

PUC-ORI"

PUC-ORI"

205 a 249 c 234 g 193 t 7 others

Query Match

Best Local Similarity

Matches

101; Conservative

0; Mismatches

32; Indels

0; Gaps

0;

Q 66 AGACTGCTCTCTCTGAATGATGGCTCTCTGATCTATCTTTGACAAAATCTACACCA 125

|||||

D 459 AGGCAGCGCTGCTCTCTGAATGATGGCTCTCTGATCTATCTTTGACAAAATCTACACCA 400

|||||

Q 126 CCAAGAGGAGCTGTGCTTACGAGATTCCTCTGGGAAATCTCTCTAGGTAAT 185

|||||

D 399 GCAGAGTGTGCTGTGCTTCTCGGAGCTCTCTCTGGGAAATCTCTCTAGGTAAT 340

|||||

Q 186 TTGGGAGAGGAA 198

|||||

D 339 GTGGCGAACGGGA 327

|||||

RESULT 8

BF087145/c

LOCUS

DEFINITION

PM1-HT0452-070900-010-c02 H10152 Homo sapiens cDNA, mRNA sequence.

623 bp mRNA linear EST 19-CCF-2000

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homalidae; Homo.

ACCESSION

BF087145

VERSION

BF087145.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homalidae; Homo.

1 (bases 1 to 623)

Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,

Nagai M.A., da Silva W. Jr., Zago M.A., Bordin S., Costa F.F.,

Goldman G.H., Carvalho A.F., Matsukuma A., Baia G.S., Simpson D.H.,

Brunstein A., de Oliveira P.S., Bucher P., Jorgensen C.V., O'Hare

M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and

Simpson A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01505-010, Sao Paulo-SP,

Brazil

Tel.: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PAPESP/LICR Human Cancer Genome

project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=812-PM1-HT0452-070

900-010-c02&t3=2000-05-07&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 18

High quality sequence stop: 575.

FEATURES

Location/Qualifiers

1..623

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0452"

/dev_stage="Adult"

/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORFES PCR (U.S. letters patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions.

BASE COUNT

155 a 162 c 132 g 163 t

ORIGIN

Query Match

Best Local Similarity

Matches

91; Conservative

0; Mismatches

22; Indels

0; Gaps

0;

Q 69 CTCGCTCTCTCTGAATGATGGCTCTCTGATCTATCTTTGACAAAATCTACACCA 128

|||||

D 460 CCGCTCTCTCTGAATGATGGCTCTCTGATCTATCTTTGACAAAATCTACACCA 401

|||||

Q 125 AGAGGAGGCTGTGCTTACGAGATTCCTCTGGGAAATCTCTCTAGGTAAT 181

|||||

D 400 AGAGTGTGCTGTGCTTCTCGGAGCTCTCTCTGGGAAATCTCTCTAGGTAAT 348

|||||

RESULT 9

BF087145/c

LOCUS

DEFINITION

PM1-HT0452-070900-010-c02 H10152 Homo sapiens cDNA, mRNA sequence.

868 bp mRNA linear EST 21-APR-2001

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homalidae; Homo.

1 (bases 1 to 868)

Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,

Nagai M.A., da Silva W. Jr., Zago M.A., Bordin S., Costa F.F.,

Goldman G.H., Carvalho A.F., Matsukuma A., Baia G.S., Simpson D.H.,

Brunstein A., de Oliveira P.S., Bucher P., Jorgensen C.V., O'Hare

M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and

Simpson A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01505-010, Sao Paulo-SP,

Brazil

Tel.: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PAPESP/LICR Human Cancer Genome

project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=812-PM1-HT0452-070

900-010-c02&t3=2000-05-07&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 18

High quality sequence stop: 575.

FEATURES

Location/Qualifiers

1..868

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0452"

/dev_stage="Adult"

/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORFES PCR (U.S. letters patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions.

BASE COUNT

155 a 162 c 132 g 163 t

ORIGIN

Query Match

Best Local Similarity

Matches

91; Conservative

0; Mismatches

22; Indels

0; Gaps

0;

Q 69 CTCGCTCTCTCTGAATGATGGCTCTCTGATCTATCTTTGACAAAATCTACACCA 128

|||||

D 460 CCGCTCTCTCTGAATGATGGCTCTCTGATCTATCTTTGACAAAATCTACACCA 401

|||||

Q 125 AGAGGAGGCTGTGCTTACGAGATTCCTCTGGGAAATCTCTCTAGGTAAT 181

|||||

D 400 AGAGTGTGCTGTGCTTCTCGGAGCTCTCTCTGGGAAATCTCTCTAGGTAAT 348

|||||

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:10:23 ; Search time 75.3682 seconds
(without alignments)
1478.802 Million cell updates/sec

Title: US-09-778-900A-3_COPY_200_500

Perfect score: 301

Sequence: 1 ggacactcgggttgatgtg.....taacaaaaactcttaagt 3.1

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 350425 seqs, 154966369 residues

total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.6	24.8	5406	10 US-09-919-408-5	Sequence 5, Appl
2	74.6	24.8	5406	10 US-09-872-136-5	Sequence 5, Appl
3	74.6	24.8	5470	10 US-09-766-078-1	Sequence 5, Appl
4	67.2	22.3	3120	10 US-09-982-610-22	Sequence 12, Appl
5	67.2	22.3	3501	10 US-09-919-408-3	Sequence 5, Appl
6	67.2	22.3	3501	10 US-09-872-136-3	Sequence 5, Appl
7	65.6	21.8	4425	10 US-09-982-610-31	Sequence 1, Appl
8	65.6	21.8	6827	10 US-09-982-610-17	Sequence 1, Appl
9	65.6	21.8	9108	10 US-09-982-610-45	Sequence 1, Appl
10	65.4	21.7	3321	10 US-09-866-510-21	Sequence 1, Appl
11	65.2	21.7	3321	10 US-09-960-352-10875	Sequence 1, Appl
12	65.2	21.7	3555	10 US-09-960-352-10875	Sequence 1, Appl
13	64.8	21.5	3270	10 US-09-866-510-1	Sequence 1, Appl
14	64.8	21.5	3270	10 US-09-866-510-3	Sequence 1, Appl
15	64.8	21.5	3270	10 US-09-866-510-5	Sequence 1, Appl
16	64.8	21.5	3270	10 US-09-866-510-7	Sequence 1, Appl
17	64.8	21.5	3270	10 US-09-866-510-9	Sequence 1, Appl
18	64.8	21.5	4054	9 US-09-955-363-35	Sequence 5, Appl
19	64.8	21.5	6378	10 US-09-919-497-40	Sequence 40, Appl

20	64.8	21.5	6412	10 US-09-766-987-1	Sequence 1, Appl
21	63.8	21.2	3321	10 US-09-866-510-13	Sequence 13, Appl
22	63.8	21.2	3321	10 US-09-866-510-15	Sequence 15, Appl
23	63.8	21.2	3321	10 US-09-866-510-17	Sequence 17, Appl
24	63.8	21.2	3321	10 US-09-866-510-19	Sequence 19, Appl
25	63.8	21.2	4465	9 US-09-955-363-1	Sequence 1, Appl
26	63.6	21.1	366	10 US-09-960-352-10578	Sequence 10878, A
27	63.6	21.1	391	10 US-09-960-352-2803	Sequence 2803, Ap
28	63.6	21.1	392	10 US-09-960-352-4293	Sequence 4293, Ap
29	63.6	21.1	402	10 US-09-960-352-14442	Sequence 14442, A
30	63.6	21.1	415	10 US-09-960-352-10423	Sequence 10423, A
31	63.6	21.1	3454	12 US-10-044-090-48	Sequence 48, Appl
32	63.6	21.1	3726	10 US-09-925-302-271	Sequence 271, App
33	60.4	20.1	410	10 US-09-960-352-4303	Sequence 4303, Ap
34	58.8	19.5	4268	10 US-09-954-456-293	Sequence 293, App
35	58.8	19.5	4268	10 US-09-954-456-1599	Sequence 1599, Ap
36	58.6	19.5	3992	10 US-09-944-807-9	Sequence 9, Appl
37	57.2	19.0	2184	10 US-09-728-952-82	Sequence 82, Appl
38	56.4	18.7	5084	9 US-09-954-531-1383	Sequence 1383, Ap
39	56.4	18.7	5084	10 US-09-967-7684-277	Sequence 277, App
40	55.6	18.5	434	10 US-09-960-352-5227	Sequence 5227, Ap
41	55.6	18.5	3453	10 US-09-919-408-1	Sequence 1, Appl
42	55.6	18.5	3453	10 US-09-872-136-1	Sequence 1, Appl
43	55.2	18.3	422	10 US-09-960-352-6956	Sequence 6956, Ap
44	54.8	18.2	2869	9 US-10-016-283-2	Sequence 2, Appl
45	53	17.6	2782	9 US-10-108-605-250	Sequence 250, App

ALIGNMENTS

RESULT 1
US-09-919-408-5
; Sequence 5, Application US/09919406
; Patent No. US2002007207A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTOPOENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/766,678
FILING DATE: 25-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,829
FILING DATE: 09-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-060
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5470 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 286..4386
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-766-678-1
Query Match 24.8%; Score 74.6; DB 10; Length 5470;
Best Local Similarity 78.8%; Pred. No. 1.8e-13;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CCGAGCTTCTCTGAATGATGCTCTGATCTATCTTGACAAATACACACCA 128
DB 3473 CCGAGCTTCTCTGAATGATGCTCTGATCTATCTTGACAAATACACCA 128
/ 129 AGAGGAGCTGTGCTTACGAGTATCTGTGGGAAATCTTCCTTAGGT 181
DB 3533 AGAGGAGCTGTGCTTACGAGTATCTGTGGGAAATCTTCCTTAGGT 181
RESULT 4
US-09-582-610-22
Sequence 22, Application US/09982610
Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
SEQUENCE: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California

COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 05/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3120 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-582-610-22
Query Match 22.3%; Score 67.2; DB 10; Length 3120;
Best Local Similarity 70.3%; Pred. No. 2.9e-11;
Matches 90; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 54 TTGCTCTCTTCTAGACTTCCTCTGAAATGATGCTCTGATCTATCTTTGACA 113
DB 2528 TTGCTAGGGCAATGCCCTCTCTCTGTAATGATGCTCTGATCTATCTTTGAG 2587
QY 114 AAATCTACAGCAACAGAGGAGCTGTGCTTACGAGTATCTGTGAGAAATCTCT 173
DB 2588 GCACTACACCAATAGAGGATGCTGCTGATGATGGAATATACIGGAGAAATCTCT 2647
QY 174 CCTTAGGT 181
DB 2648 CAGTTGGT 2655
RESULT 5
US-09-919-408-3
Sequence 3, Application US/09919408
Patent No. US2002007207A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ilmor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408

Matches 90; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 54 TTGCGCTCTTACAGTCGACTCTCTCGAAATGAGGCTCTCTGATCTATCTTIGACA 113
|||||
Db 2585 TTGTACGGGCAATGGCCGCTCTCCCTGTAAATGATGGCCCGGAAAGCGTGTGAAG 2644
|||||
QY 114 AATCTACAGCAGCAAGGAGCGCTGTCTTACGGAGTATGCTGTGGGAAATCTTCT 173
|||||
Db 2645 GCATCTACACATTAAGAGTATGCTGTCTATATGGAATATTACTGTGGGAAATCTTCT 2704
|||||
QY 174 CTTAGGT 181
|
Db 2705 CACTTGGT 2712
|

RESULT 7
US-09-982-610-31
; Sequence 31, Application US/09982610
; Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982.610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446.648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40.378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 4425 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-982-610-31

Query Match 21.8%; Score 65.6; DB 10; Length 4425;
Best Local Similarity 74.1%; Pred. No. 1.1e-10;
Matches 83; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 69 CTCGACTCTCTCGAAATGAGGCTCTCTGATCTATCTTIGACAAATCTACACACCA 128
|||||

Db 2553 CCCGCTCCCCCTGAGTGGAGGCTCTCTGATCTATCTTIGACAAATCTACACACCA 2494
|||||

QY 129 AGAGCGAGCTGTCTTACGAGTATGCTGTGGGAAATCTTCTTAGG 180
|||||

Db 2493 AGAGTGAAGTGTCTTGTGGGCTCTCTGAGATCTTCTCTGGG 2442
|||||

RESULT 9

Db 3251 CCCGCTCCCCCTGAGTGGAGGCTCTCTGATCTATCTTIGACAAATCTACACACCA 3310
|||||

QY 129 AGAGCGAGCTGTCTTACGAGTATGCTGTGGGAAATCTTCTTAGG 180
|||||

Db 3311 AGAGTGAAGTGTCTTGTGGGCTCTCTGAGATCTTCTCTGGG 3362
|||||

RESULT 8
US-09-982-610-17/c
; Sequence 17, Application US/09982610
; Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982.610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446.648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40.378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 6827 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-982-610-17

Query Match 21.8%; Score 65.6; DB 10; Length 6827;
Best Local Similarity 74.1%; Pred. No. 1.3e-10;
Matches 83; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 69 CTCGACTCTCTCGAAATGAGGCTCTCTGATCTATCTTIGACAAATCTACACACCA 128
|||||

Db 2553 CCCGCTCCCCCTGAGTGGAGGCTCTCTGATCTATCTTIGACAAATCTACACACCA 2494
|||||

QY 129 AGAGCGAGCTGTCTTACGAGTATGCTGTGGGAAATCTTCTTAGG 180
|||||

Db 2493 AGAGTGAAGTGTCTTGTGGGCTCTCTGAGATCTTCTCTGGG 2442
|||||


```

1  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
2
3  FILE REFERENCE: EM-104.01
4
5  CURRENT APPLICATION NUMBER: US/09/866,510
6
7  CURRENT FILING DATE: 2001-05-25
8
9  PRIOR APPLICATION NUMBER: 60/250,747
10
11 PRIOR FILING DATE: 2000-12-01
12
13 PRIOR APPLICATION NUMBER: 60/289,103
14
15 PRIOR FILING DATE: 2001-05-07
16
17 NUMBER OF SEQ ID NOS: 33
18
19 SOFTWARE: PatentIn ver. 2.1
20
21 SEQ ID NO 3
22
23 LENGTH: 3270
24
25 TYPE: DNA
26
27 ORGANISM: Homo sapiens
28
29 FEATURE:
30
31 NAME/KEY: CDS
32
33 LOCATION: (1)..(3267)
34
35 PS-09-866-510-1

```

Query Match	21.5%	Score 64.8	DB 10	Length 3270
Best local Similarity	75.0%	Pred. No. 1.7e-10		
Matches 81	Conservative 0	Mismatches 27	Indels 0	Gaps 0
QY 74	CTTCTCTGAAATGATGGCTCTCTGAATCTATCTTTGACAAATCTTACAGACACAGAGC	133		
DD 2565	CTGCCCGTGAAGGATGGCTCTCTGAGAGCATCTTTGACAACCTCTACACCACACTGAGT	2628		
QY 134	GACGTGTGGTCTTTACGGAGTATTCGTCTGGGAATATCTTCTCTTACGT	181		
DD 2629	GATCTCTGGTCTTAAGGCAATTCGTCTCTGGAGATCTTTTCCCTTGGT	2676		

```

RESULT 15
US-09-866-510-5
: Sequence 5, Application US/09866510
: Patent No. US20020111304A1
: GENERAL INFORMATION:
: APPLICANT: KAZLAUSKAS, ANDRIUS
: APPLICANT: IKUNO, YASUSHI
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
: FILE REFERENCE: BEM-104.01
: CURRENT APPLICATION NUMBER: US/09/866,510
: CURRENT FILING DATE: 2001-05-25
: PRIOR APPLICATION NUMBER: 60/250,747
: PRIOR FILING DATE: 2000-12-01
: PRIOR APPLICATION NUMBER: 60/289,103
: PRIOR FILING DATE: 2001-05-07
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 5
: LENGTH: 3270
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(3267)
US-09-866-510-5

```

Query Match	21.5%	Score 64.8	DB 10	Length 3270
Best Local Similarity	75.0%	Prsd. No. 1.7e-10		
Matches	81	Conservative 0	Mismatches 27	Indels 0
Gaps	0			
QY	74	CTTCTCTGGAATGGATGGCTCTGTAATCTATCTTTGACAAAATCTACACACCAAGAC	133	
DB	268	CTGGCCGTAAGTGGATGGCTCTGACAGCATCTTTGACAACTCTACACCACTGACT	2628	
QY	134	GACGCTGGCTCTTACGGAGTATTGCTGTGGAAATCTTCTCTTAGG	181	
DB	2629	GATCTCTGGTCTTATGGCAATCTCTCTCTGGGAGATCTTTCCCTTGGT	2676	

Search completed: December 8, 2002, 09:25:43
Job time : 98.3682 secs


```

repeat_region 32425..32610
/Note="MERSA repeat: matches 4..189 of consensus"
repeat_region 32946..33076
/Note="FLAM_A repeat: matches 1..133 of consensus"
repeat_region 33231..33323
/Note="MIR repeat: matches 118..222 of consensus"
repeat_region 33433..33747
/Note="AluYb8 repeat: matches 1..313 of consensus"
repeat_region 35242..35605
/Note="THE1B repeat: matches 1..364 of consensus"
repeat_region 35886..36063
/Note="89 copies 2 mer tt 57% conserved"
repeat_region 36786..36959
/Note="MER20 repeat: matches 2..169 of consensus"
repeat_region 37212..37318
/Note="MIR repeat: matches 131..241 of consensus"
repeat_region 37283..37330
/Note="L2 repeat: matches 2688..2739 of consensus"
repeat_region 37659..37781
/Note="FLAM_C repeat: matches 1..125 of consensus"
repeat_region 38866..39047
/Note="MERSA repeat: matches 1..186 of consensus"
repeat_region 39114..39348
/Note="MIR repeat: matches 4..231 of consensus"
repeat_region 40010..40203
/Note="AluJ0 repeat: matches 86..274 of consensus"
repeat_region 40286..40581
/Note="AluSFP repeat: matches 1..297 of consensus"
repeat_region 41033..41330
/Note="AluSFP repeat: matches 1..299 of consensus"
repeat_region 43786..43971
/Note="MERSB repeat: matches 1..177 of consensus"
repeat_region 44671..44839
/Note="MERSA repeat: matches 8..189 of consensus"
repeat_region 45327..45508
/Note="MERS repeat: matches 1..193 of consensus"
repeat_region 46166..46290
/Note="LIMC5 repeat: matches 7796..7925 of consensus"
repeat_region 46331..46510
/Note="MIR repeat: matches 74..256 of consensus"
repeat_region 46598..46767

```

```

Query Match 99.5% Score 299.4; DB 9; Length 164519;
Best Local Similarity 99.7%; Pred. No. 3.3e-78;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACACTCGGTTGAATGCTTTGTTTCATGCTCTCTCTCAAGGCCCTTGGCCCT 60
b 87314 GGACACTCGGTTGAATGCTTTGTTTCATGCTCTCTCTCAAGGCCCTTGGCCCT 87255

QY 61 CTCTAGACTCGACTTCTCTGAATGGATGGCTCTCTGAATCTATCTTTGACAAATCTA 120
Db 87254 CTCTAGACTCGACTTCTCTGAATGGATGGCTCTCTGAATCTATCTTTGACAAATCTA 87195

QY 121 CAGCACCAGACGAGCTGTGCTTACGAGTATGCTGTGGAAATCTTCTCTTAGG 180
Db 87194 CAGCACCAGACGAGCTGTGCTTACGAGTATGCTGTGGAAATCTTCTCTTAGG 87135

QY 181 TAAATTTGGAGAGGAGAAATCAACAGCCCAAGAAATTAATGCTGTGATCTTGTG 240
Db 87134 TAAATTTGGAGAGGAGAAATCAACAGCCCAAGAAATTAATGCTGTGATCTTGTG 87075

QY 241 AATGTCCTTTGGTGGACGCTTTAGATTAGAACTACTGTACAAAATCTTAAAG 300
Db 87074 AATGTCCTTTGGTGGACGCTTTAGATTAGAACTACTGTACAAAATCTTAAAG 87015

QY 301 T 301
Db 87014 T 87014

```

RESULT 2
AC128808/c

LOCUS DEFINITION

AC128808 101987 bp DNA linear HTG 24-JUL-2002
Rattus norvegicus clone CH230-69N22, *** SEQUENCING IN PROGRESS
***, 59 unordered pieces.

ACCESSION VERSION

AC128808.1 GI:21952667
HTG: HGSC-PHASE1.

KEYWORDS SOURCE

Rattus norvegicus.
Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE AUTHORS

1 (bases 1 to 101987)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Alsbrooks, S.H., Amarantunga, H.C., Are, J.R., Ayala, M., Banks, T.,
Bartolacci, J., Benton, J., Blum, K., Blankenburg, K., Blom, D.,
Bouchard, J., Bowie, S., Briveau, M., Brown, E., Brown, M., Bryant, N.P.,
Burch, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, F.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Choudhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Deinewate, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, I., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorelli, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Herrandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Homi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, I., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, I.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louissegh, H.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mathew, P., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Mosser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nockemwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tanerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

TITLE JOURNAL

Unpublished
2 (bases 1 to 101987)
Worley, K.C.

REFERENCE AUTHORS

Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

TITLE JOURNAL

Unpublished
2 (bases 1 to 101987)
Worley, K.C.

COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

PROJECT INFORMATION

Center project name: KAMN
Center clone name: CH230-69N22

SEQUENCING VECTOR

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 55655 bases at least Q40

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Brooks, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 4597 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-175

Query Match 27.5%; Score 27.8; DB 4; Length 4597;
Best Local Similarity 59.5%; Pred. No. 7.2;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 23 ACTCATAGAGCTGGATGCGCTTTCCTGCTTGATTAATCTTCCTGCTTCTTCATAGAT 82
DB 3778 AATTAATACTCTGTGAACCGTTTATTTATGATAATTCATCTATTTGCAAGTACT 3719

QY 83 ATGTAAATGCTTCAAGTT 101
DB 3718 AATAAATTTTATTT 3700

RESULT 5
US-09-419-568F-25
Sequence 25, Application US/09419568F
Patent No. 6331613
GENERAL INFORMATION:
APPLICANT: Dumoutier, Laure
APPLICANT: Louhed, Jamila
APPLICANT: Renauld, Jean-Christophe
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode I Cell Inducible Fac
FILE REFERENCE: LUD 5543.2
CURRENT APPLICATION NUMBER: US/09/419,568F
CURRENT FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: US09/354,243
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: US09/178,973
PRIOR FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 29
SEQ ID NO 25
LENGTH: 4757
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-419-568F-25

Query Match 27.5%; Score 27.8; DB 4; Length 4797;
Best Local Similarity 65.1%; Pred. No. 7.3;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 39 TGCCTTTCCTGCTTTCATTAATCTTCCTGCTTCTTCATAGATACTAAATGCTTCAAA 98
DB 3056 TCCCGAGCACTAATAAGAGTGTGACTGTTCTCTCTTGTATATAAATGAAGGCTTTGTA 3117

QY 59 GTT 101
DB 3118 GTT 3120

RESULT 6
US-09-354-243B-25
Sequence 25, Application US/09354243B
Patent No. 6359117
GENERAL INFORMATION:
APPLICANT: Dumoutier, Laure
APPLICANT: Louhed, Jamila
APPLICANT: Renauld, Jean-Christophe
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode I Cell Inducibl
FILE REFERENCE: LUD 5543.1
CURRENT APPLICATION NUMBER: US/09/354,243B
CURRENT FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: US09/178,973
PRIOR FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 29
SEQ ID NO 25
LENGTH: 4797
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
ATTORNEY/AGENT INFORMATION:

Query Match 27.5%; Score 27.8; DB 4; Length 4797;
Best Local Similarity 65.1%; Pred. No. 7.3;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 39 TGCCTTTCCTGCTTTCATTAATCTTCCTGCTTCTTCATAGATACTAAATGCTTCAAA 98
DB 3056 TCCCGAGCACTAATAAGAGTGTGACTGTTCTCTCTTGTATATAAATGAAGGCTTTGTA 3117

QY 59 GTT 101
DB 3118 GTT 3120

RESULT 7
US-08-144-602B-4
Sequence 4, Application US/08144602
Patent No. 5641876
GENERAL INFORMATION:
APPLICANT: McElroy, David
APPLICANT: Wu, J
TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIX, HARGRAVE, DEVANS & DOYLE LLP
5 ALLEY CLINTON SQUARE, P.O. BOX 1051
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/144,602B
FILING DATE: 27-OCT-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:


```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/082,310
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0520 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4156
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSINOT07
; CLONE: 2124957
; -09-575-205-3

```

```

Query Match      26.5%; Score 26.8; DB 4; Length 1894;
Best Local Similarity 64.5%; Pred. No. 12;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 20 CTGACTCATTAGAGCTGGATGCGTTCCTGCTGCTGATGATTAATTCCTGTTCTTCATTA 79
Db 1466 CTTAATTTTCATTCCTGGTGGATTTCTCTCTGGCAATTCCTTTTGTACCCACTA 1407

Qy 80 GA 81
Db 1406 AA 1405

```

```

RESULT 11
US-09-457-066-42
; Sequence 42, Application US/09457066
; Patent No. 6432673
; GENERAL INFORMATION:
; APPLICANT: Gao, Zeren
; APPLICANT: Hart, Charles E.
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Gilbertson, Debra G.
; APPLICANT: West, James W.
; TITLE OF INVENTION: GROWTH FACTOR HOMOLOG ZVEGF3
; FILE REFERENCE: 98-60
; CURRENT APPLICATION NUMBER: US/09/457.066
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 3571
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1049)...(2086)
US-09-457-066-42

```

```

Query Match      26.1%; Score 26.4; DB 4; Length 3573;
Best Local Similarity 69.2%; Pred. No. 18;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 37 GATCCCTTTCTGCTGATGATTCCTGCTGCTGATGATTAATTCCTGTTCTTCATTAAGTAA 88
Db 2813 GTTGGCTTTCTAATCTGTTAAATATTCCTATTTTACCAAGGATTTAA 2864

RESULT 12

```

```

US-09-302-620B-90/C
; Sequence 90, Application US/09302620B
; Patent No. 6331420
; GENERAL INFORMATION:
; APPLICANT: Wilson, C. Ron
; APPLICANT: Craft, David L.
; APPLICANT: Ehrich, Dudley
; APPLICANT: Esboo, Mark
; APPLICANT: Madduri, Krishna M.
; APPLICANT: Cornett, Cathy A.
; APPLICANT: Brenner, Alfred A.
; APPLICANT: Tang, Maria
; APPLICANT: Loper, John C.
; APPLICANT: Gleeson, Martin
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450
; TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA
; TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS
; FILE REFERENCE: 1010-16.seq
; CURRENT APPLICATION NUMBER: US/09/302.620B
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 90
; LENGTH: 3826
; TYPE: DNA
; ORGANISM: Candida tropicalis
US-09-302-620B-90

```

```

Query Match      26.1%; Score 26.4; DB 4; Length 3826;
Best Local Similarity 69.2%; Pred. No. 19;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 25 TCATGAGCTGGATGCTTTCTCTGCTGCTGATGATTAATTCCTGTTCTTCATTA 76
Db 3291 TCATGAGCTGGATGCTTTCTCTGCTGCTGATGATTAATTCCTGCTGCTTCATTA 3240

```

```

RESULT 13
US-08-477-451-9/C
; Sequence 9, Application US/08477451
; Patent No. 5928865
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; TITLE OF INVENTION: Helicobacter Pylori CagI Region
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chilton Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477.451
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0335.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5599 base pairs
; TYPE: nucleic acid

```


PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer

XX Claim 1: SEQ ID No 73; 27pp; English.

XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 438 and 500 are missing from the sequence listing) sequences
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC peptide nucleic acid-oligonucleotide (PNA) or at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters for the diagnosis
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC http://wipo.int/pub/published_pct_sequences.

XX Sequence 6316 BP; 1490 A; 255 C; 1803 G; 2768 T; 0 other;

Query Match 65.1%; Score 613; DB 22; Length 6316;
Best Local Similarity 78.2%; Pred. No. 6.4e-114;
Matches 736; Conservative 0; Mismatches 205; Indels 0; Gaps 0;
QY 1 GCGGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGAGCTTCCTCCAAAG 60
DB 4272 GTCGAAAGATACGATAGCTTTTGGGATTGAGTTGTTGCTAGTTTITTAAG 4351
QY 61 GTGGCAACAGAGCTAGTTCCTCCAGGCTCCAGGTCAGTTCGCTGCGGAGGT 120
DB 4332 GTGTTAAGTAAGCTTGTAGTTTITTAGCGCTTTAGTTAGTTTGTGTCGAGGGI 4351
121 CTCGGTCCCTCTCTAGACTCTCGGACAGCTCAAGGGGTCAGGAGCGGGGACAGC 180
4392 TTTCGGGTTTTTATGATTTTCGGGATAGTTTGAAGGGTTAGGAGCGCGGATAGC 4451
181 GCGGAAGACAGGACAGGAGGACAGGCTCGGCTCAGTCTCGTGCCTCAAGAAC 240
4452 GCGGAAGATAGTAGTAAGGAGGATAGTCTGATTCGTTTTCGTGTAAGAAAT 4511
QY 241 ACCGTCGGGAGGCGGCGGACAGCTCCCTGATCGGACTTTCGCGCTAGGCGCAGC 300
4512 ATGTCGCGGAGGCGGCTAGTTTTCGATCGGATTTTCGTTTTCGTTTTCGTTTTCG 4571
QY 301 GCGGAGCTTCAGCTTCTCCCTCCAGTTTCGCGGCGGCGGCGGCTAGTAAAGC 360
4572 GCGGAGTTTATGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCG 4631
QY 361 GGTGTAGGAGGATCGCAAGATTTCCTGAGCGGATGGCAGGAGGAGGAGGCA 420
4632 GGTGTAGGAGGATGTAAGGATTTTCAGCGCGATGGTGTAGGAGGAGGATAGGTA 4631
QY 421 AGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
4692 AGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4751
QY 481 AGCAATCCGACGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 540
4752 AGTAATTTTACGCGGCTCGGCTTCGGTTTCGGTTTTCGTTTTCGTTTTCGTTTTCG 4811

QY 541 CGTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 4812 TGTATGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4871
QY 601 GGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 4872 GGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4931
QY 661 AAGGTATATAAT 720
DB 4932 AAGGTATATAAT 4991
QY 721 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 4992 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5051
QY 781 CTCGGGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 840
DB 5052 TTTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5111
QY 841 GGTGTATCT 900
DB 5112 GGTGTAT 5171
QY 901 ACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 941
DB 5172 ACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5212

RESULT 4

ABL70215

ID ABL70215 standard; DNA; 6316 BP.

XX ABL70215;

XX ABL70215;

XX 01-JUL-2002 (first entry)

XX Chemically treated cell signalling DNA sequences53.

XX Cell signalling; cytosine methylation; cell signalling disease;

XX cancer; tumour; cytostatic; ds.

XX Unidentified.

XX W0200202807-42.

XX 10-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07471.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A. Piepenbrock C. Berlin K.

XX WPI; 2002-154758/20.

XX The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA or genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligonucleotides for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and

DL 1144 CGAATAAAGCACTTAACGACCGAATCGTTAACCGCG 1107

RESULT 10

AA561171/C

ID AA561171 standard; DNA: 6316 BP.

XX AA561171;

XX 29-JAN-2002 (first entry)

XX Human gene regulation-associated gene oligonucleotide #126.

XX Human: Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preseclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; cesophagial cancer; ds; tumour;
 immunostimulant; cardiac; antinflammatory; coagulant; antiasthmatic;
 nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

XX Homo sapiens.

XX W0200177375-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP03968.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI: 2002-017470/02.

XX New nucleic acid sequences from chemically modified genes associated
 with gene regulation, useful for analysing cytosine methylations for
 diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 disease

XX Claim 1: SEQ ID NO 130; 26pp; English.

The invention relates to 224 nucleic acid sequences comprising at least
 18 bases of a chemically pretreated gene associated with gene regulation
 selected from 43 known genes (or complementary sequences). The
 chemical pretreatment converts cytosine bases unmethylated at the
 5-position to uracil or another base with hybridisation behaviour
 dissimilar to cytosine, to enable analysis of cytosine methylations.
 The DNA sequences, oligomers (or sets/arrays) and method are
 useful in the diagnosis of diseases (or predisposition to diseases)
 associated with gene regulation and in therapy of such diseases, by
 enabling analysis of the cytosine methylation patterns of such genes,
 kits are provided. They are especially useful in diagnosis
 and therapy of e.g. severe combined immunodeficiency disease, cardiac
 disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 preseclampsia, graft versus-host disease. The present sequence is a
 sequence included in the sequence data for this specification and is
 associated with the human gene regulation-associated genes.
 Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic
 format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

XX Sequence 6316 BP; 1474 A; 255 C; 1549 G; 3038 T; 0 other;

XX Query Match 54.3%; Score 510.8; DB 24; Length 6316;

XX Best Local Similarity 71.5%; Pred. No. 1.8e-93;

	Matches 671;	Conservative	0;	Mismatches 267;	Indels	0;	Gaps	0;
QY 2	CGAAAAAAGACAGGACAGCTCCCTGGACCTGAGCTGGTTCGAGCTTCCTCCAAAGG	61						
DB 2044	CGAAAAAAGACAGGACAGCTCCCTAAACCTAAATTCGCAATTCCTCCCAAAA	1985						
QY 62	TCCCAAGCAACGCTCAGTTCCTCTAGGCGCTCCAGGTCAGTTCCTTCGCGAGGTC	121						
DB 1984	TACCAACAAGCTCAATTCCTCTAAAGCTTCAATTCCTTCCTTCCTTCCTTCCTTC	1925						
QY 122	TCCGCTGCTTCCTAGACCTCTCGGACAGCTCTGAGGCTCAGGAGCGGAGACAGCG	181						
DB 1924	TCGATACCTTCCTTAACCTTCGAAACAATCTAAAAAATCAAAAACGACGAACAGC	1865						
QY 182	CGGAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	241						
DB 1864	CGAAAAAAGCAACAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1805						
QY 242	CGCTCGGCGAGCGCGCGGAGCTTCCTTCGATCGGCTTCGCGCTTCGCGCTTCGCG	301						
DB 1804	CGCTCGGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1745						
QY 302	GGGAGCTTCAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC	361						
DB 1744	ACGAAACCTCAACCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT	1685						
QY 362	GGTGGAGGAGCTTCGAAAGGATTCCTGAGCGGATGGGAGGAGGAGGAGGAGGAGGAGG	421						
DB 1684	AATAAAAAAATCTACAAAAATTCCTTAAGCGGATTAACAAAAAAGGAGGAGGAGGAGG	1625						
QY 422	GAGGCGCGGAGCAAGACACCTTGAACCTTCGCGCGCGCGCTCCCGCGCGCGCGCTCC	481						
DB 1624	AAAAAGCGGAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1565						
QY 482	GCACCTCCCGACGCGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC	541						
DB 1564	ACACCTCCCGACGCGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC	1505						
QY 542	CGTACCGCGAGGAGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	601						
DB 1504	CGTACCGCGAGGAGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1445						
QY 602	GGTGGGCGACCTTCAGCTTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	661						
DB 1444	AATAAAAAAATCTAGCTACCAAAAAAATACCAAAAAAATACCAAAAAAATACCAAAAA	1385						
QY 662	AGGTATTAATCGCTC	721						
DB 1384	AAATTAATTAATCGCTC	1325						
QY 722	CG	781						
DB 1324	CG	1265						
QY 782	TCCGCGCGCTGGGTCAGCG	841						
DB 1264	TCCGCGCGCTGGGTCAGCG	1205						
QY 842	GTTCCTTCCTTCGCTGGGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	901						
DB 1204	ATATCTTCCTTCGCTGGGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1145						
QY 902	CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	939						
DB 1144	CGAAAAAAGCACTTCCTTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1107						

RESULT 11

ABK31261/C

ID ABK31261 standard; DNA: 6316 BP.

XX AC

XX ABK31261;

XX XX

PF 06-DEC-2001; 2001WO-EP14320.
 XX PR 06-DEC-2000; 2000DE-1061338.
 XX PA (EPIG-) EPIGENOMICS AG.
 XX PT Schacht O;
 XX WPI; 2002-500450/53.
 XX New nucleic acid fragments from chemically treated
 XX angiogenesis-associated genes, useful for determining methylation
 XX status, e.g. in diagnosis or treatment of cancer
 XX Claim 1; SEQ ID NO 94; 41pp + Sequence Listing; German.
 XX the invention relates to a nucleic acid (I) comprising a segment of 16
 XX bases of chemically pretreated DNA of angiogenesis-associated genes (II)
 XX having sequences (AB066971-AB067178) or their complements (I), also
 XX related oligomers, are used to evaluate the methylation status and/or
 XX single-nucleotide polymorphisms, in angiogenesis-related genes, for
 XX diagnosis and treatment of eye diseases, proliferative retinopathy,
 XX neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
 XX diabetic retinopathy, macular degeneration caused by neovascularisation,
 XX psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
 XX Crohn's disease.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 6321 BP: 1474 A; 258 C; 1553 G; 3036 T; 0 other;
 Query Match 54.3%; Score 510.8; DB 24; Length 6321;
 Best Local Similarity 71.5%; Pred. No. 1.8e-33;
 Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;
 QY 2 CGGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 61
 Db CGGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 1990
 QY 62 TCCGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 121
 Db TCCGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 1930
 QY 122 TCCGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 181
 Db TCCGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 1870
 QY 182 CGGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 241
 Db CGGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 1810
 QY 242 CGGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 301
 Db CGGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 1750
 QY 302 CGGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 361
 Db CGGAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGCGAGTCTCCCAAGG 1690
 QY 362 GGTGAGGAGCTCTGAGAGGATTCCTGAGGCGATGGGAGGAGGAGGAGGAGGAA 421
 Db GGTGAGGAGCTCTGAGAGGATTCCTGAGGCGATGGGAGGAGGAGGAGGAGGAA 1630
 QY 422 GAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
 Db GAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1570
 QY 482 GACACTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 541
 Db GACACTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1510
 QY 542 CGTAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601

Db 1509 CTTAACCCTGCAAAACAGACCTTAAATAAATAAATAAATAAATAAATAAATAA 1450
 QY 602 GGTGGGGGACCTGTGACGTACACAGAGAGGTGCCGGGTAGGAGTGGCTGGGAA 661
 Db 1449 AATAAAGACCTTAACTGACCAAAATAAATAAATAAATAAATAAATAAATAA 1390
 QY 662 AGTTTATAAAAGCCCTGCGCTGCTTCTATCGAGTCCGCGAGGCTGGGAG 721
 Db 1389 AATAAATAAAAGCCCTGCGCTGCTTCTATCGAGTCCGCGAGGCTGGGAG 1330
 QY 722 CGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 781
 Db 1329 CGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1270
 QY 782 TCCGGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
 Db 1269 TCCGAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1210
 QY 842 GTGTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 901
 Db 1209 AATAAAGACCTTAACTGACCAAAATAAATAAATAAATAAATAAATAAATAA 1150
 QY 902 CGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 939
 Db 1149 CGAATAAAGACCTTAACTGACCAAAATAAATAAATAAATAAATAAATAAATAA 1112
 RESULT 13
 AAV19347
 ID AAV19347 standard; DNA; 2313 BP.
 XX AC AAV19347;
 XX DI 17-AUG-1998 (first entry)
 XX DE Human soluble VEGF receptor FLT-1 cDNA.
 XX KW FLT-1; sFLT-1; vascular endothelial growth factor; VEGF; sVEGF-R1;
 XX KW receptor tyrosine kinase; signal transduction; angiogenesis;
 XX KW gene therapy; tumour; metastasis; inflammation; psoriasis;
 XX KW rheumatoid arthritis; haemangioma; diabetic retinopathy;
 XX KW angiofibroma; macular degeneration; human; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FI CDS 250..2313
 XX FI /*tag= a
 XX PN W09813071-A1.
 XX PD 02-APR-1998.
 XX PF 24-SEP-1997; 87WO-US17044.
 XX PR 24-SEP-1996; 86US-0026641.
 XX PA (MERI) MERCK & CO INC.
 XX PI Bett A, Goldman CK, Huckle WR, Kendall RL, Thomas KA;
 XX DR WPI: 1998-23042/20.
 XX DR P-PSDS: AAW4484.
 XX DR Inhibition of angiogenesis, particularly in tumours - by using DNA
 XX PI expressing a soluble form of a tyrosine kinase receptor which forms
 XX PI a dimer with a vascular endothelial growth factor
 XX PS Example 1; Fig 1: 52pp; English.
 XX CC This cDNA clone codes for human soluble FLT-1 (sFLT-1, see AAW4484),
 XX CC the tyrosine kinase receptor (TKR) for human vascular endothelial

CC growth factor (VEGF). PCR cloning was used to obtain the flt-1
 CC derived TKR (sVEGF-RI) flt-1 cDNA from a HUVEC cDNA library. The
 CC invention relates to methods of gene therapy for inhibiting
 CC angiogenesis associated with solid tumor growth, tumour metastasis,
 CC psoriasis, rheumatoid arthritis, haemangioma, diabetic retinopathy,
 CC angiofibroma and macular degeneration. For primary tumour growth
 CC and metastasis, this involves transfer of a nucleotide sequence
 CC encoding a soluble form of VEGF TKR to a mammalian (human) host.
 CC preferably using an adenovirus or recombinant plasmid DNA vector.
 CC The transferred nucleotide sequence encodes mRNA and a soluble
 CC receptor protein which binds to VEGF in extracellular regions
 CC adjacent to the primary tumor and vascular endothelial cells.
 CC Formation of a sVEGF-R/VEGF complex prevents binding of VEGF to the
 CC FLT-1 TKR, antagonising transduction of the normal intracellular
 CC signals associated with vascular endothelial cell-induced tumour
 CC angiogenesis. Expression of a soluble TKR may also impart a
 CC therapeutic effect by binding either with or without VEGFs to form
 CC non-functional heterodimers with full-length VEGF-specific TKRs and
 CC thereby inhibiting the mitogenic and angiogenic activities of
 CC VEGFs.

-J Sequence 2313 BP; 705 A; 537 C; 539 G; 532 T; 0 other;

Query Match 20.7%; Score 194.8; DB 19; Length 2313;
 Best Local Similarity 98.1%; Pred. No. 3.2e-30;
 Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 730 GGGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGGAGCGGCTCGGGGC 789
 Db 1 GCGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGGAGCGGCTCGGGGC 60
 QY 790 TCGGCTGAGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 849
 Db 61 TCGGCTGAGCGGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 118
 QY 850 CTGGCTGAGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 849
 Db 119 CTGGCTGAGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 118
 QY 910 CGGACTCTGCGGCGGCGGCGGCTCTTTGGCGCGGG 941
 Db 179 CGGACTCTGCGGCGGCGGCGGCTCTTTGGCGCGGG 210

RESULT 14
 AAV09330
 ID AAV09330 standard; cDNA; 2313 BP.

AAV09330;

13-MAY-1998 (first entry)

Soluble VEGF receptor/inhibitor (sVEGF-RI) encoding cDNA.

Vascular endothelial cell growth factor; VEGF; receptor; inhibitor;
 soluble; treatment; angiogenesis; psoriasis; tumour; arthritis; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 250..2313

/*tag= a

/product= "soluble VEGF inhibitor"

US5712380-A.

27-JAN-1998.

21-APR-1994; 94US-0232538.

21-APR-1994; 94US-0232538.

25-MAR-1993; 93US-0038769.

PA (MERI) MERCK & CO INC.
 XX Kendall RL, Thomas KA;
 XX WPI; 1998-120332/11.
 DR P-PSDB; AAW47036.
 XX Nucleic acid encoding soluble form of vascular endothelial cell
 PI growth factor; receptor - and related vector and transformed cells,
 PI expressing soluble inhibitor of VEGF useful for inhibiting
 PI angiogenesis. e.g. for treatment of psoriasis, arthritis, tumours
 etc.

Claim 3; Fig 2A-B; 47pp; English.

CC this cDNA encodes a soluble inhibitor of vascular endothelial cell growth
 CC factor (VEGF). The VEGF receptor/inhibitor (sVEGF-RI) represents a
 CC fragment of the VEGF receptor that binds VEGF with high affinity but is
 CC unable to transduce a signal. sVEGF-RI is used to inhibit VEGF activity,
 CC specifically VEGF-induced angiogenesis, particularly for treatment of
 CC psoriasis, rheumatoid arthritis, haemangiomas, angiofibromas, diabetic
 CC retinopathy, neovascular glaucoma or tumour vascularisation. The sVEGF-RI
 CC can be administered topically or intravenously or from slow-release
 CC matrices.

Sequence 2313 BP; 705 A; 537 C; 539 G; 532 T; 0 other;

Query Match 20.7%; Score 194.8; DB 19; Length 2313;
 Best Local Similarity 98.1%; Pred. No. 3.2e-30;
 Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 730 GGGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGGAGCGGCTCGGGGC 789
 Db 1 GCGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGGAGCGGCTCGGGGC 60
 QY 790 TCGGCTGAGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 849
 Db 61 TCGGCTGAGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 118
 QY 850 CTGGCTGAGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 849
 Db 119 CTGGCTGAGCGGCGGCGGCGGCTCGGCGGAGGATTACCGGGGAGAGTGGTGTCTC 118
 QY 910 CGGACTCTGCGGCGGCGGCGGCTCTTTGGCGCGGG 941
 Db 179 CGGACTCTGCGGCGGCGGCGGCTCTTTGGCGCGGG 210

RESULT 15
 AAV01457
 ID AAV01457 standard; cDNA; 2523 BP.

AAV01457;

25-MAR-1998 (first entry)

Human VEGF receptor extracellular domain coding sequence.

Immunoglobulin-like domain; extracellular domain; VEGF; human; primer;
 Vascular endothelial cell growth factor receptor; amplification; PCR;
 truncation; mutant; tumour; vascularisation; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 250..2523

/*tag= a

/product= "VEGF receptor extracellular domain"

/note= "no stop codon is given at the 3' end of the

sequence"

sig_peptide 250..315

/*tag= b

mat_peptide 315..2523


```
Db 5172 ACAGAGACGGAATTTGGCGGTGGGTCTGTTGGTGGCGGG 5212
RESULT 10
AX251106/c
LOCUS AX251106 6316 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 74 from Patent WO0168912.
ACCESSION AX251106
VERSION AX251106.1 GI:15984529
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 6316)
AUTHORS Olek.A., Piepenbrock.C. and Berlin.K.
TITLE Diagnosis of diseases associated with tumor suppressor genes and
        oncogenes
JOURNAL Patent: WO 0168912-A 74 20-SEP-2001;
        Epigenomics AG (DE)
FEATURES
    source
        1..6316
            Location/Qualifiers
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1474 a 255 c 1549 g 3038 t
ORIGIN
Query Match 54.3%; Score 510.8; DB 6; Length 6316;
Best Local Similarity 71.5%; Pred. No. 6.8e-74;
Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;
Y 2 CGGAAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCGAGCTTCCTAAAGG 61
D 2044 CGGAAAAGACGACGACGCTCCCTAAACCTTAACCTTAATTCGCAATCTCCCAAA 1985
Y 62 TCCGAAAGACGCTGAGTTCCTCCCTGAGGCTTCCAGTTCAGTTCCTGCTGGCGG 121
D 1984 TACCAACAAACGCTCAATTCCTCCCTCAACGCTTCAATTAATTAACCTTAAC 1925
Y 122 TCCGCTGCTTCCTAGACTTCTCGGACAGCTCTGAAGGGTCTAGGAGCGCGGAC 181
D 1924 TCCGATACCTTCTTAACCTTCTCGAACAATCTAATAAATCAAAACGACGACAC 1865
Y 182 CGGGAAGACGACGAGGAGACGCGGAGCTCGGCTCAGTTCCTTCCTGCGGAGG 241
D 1864 CGAATAAACAACGACGACGCTCCCTAAACGCTTCAATTAATTAACCTTAAC 1805
Y 242 CGGTGGGAGCGCGGCGGAGCTTCCTTGGATCGGACTTTCGCGCTTAGGCGGCG 301
D 1804 CCGTCGGAAAACGCGACCACTTCCTTAATTCGAACCTTCCGCGCTTAAACCA 1745
Y 302 GCGGAGCTTCAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 361
D 1744 ACGAACTTCAACCTTATCCCTTCCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1685
Y 362 GTTGAGGAGGTCTGCAAGGATTCCTGAGGCGATGGGAGGAGGAGGCAACGCA 421
D 1684 AATAAATAAATACTACAAAATTTCTTAACGCGATTAACGCGATTAACGCG 1625
Y 422 GAGGCGCGGAGGAGCAACCTTCGACTTCGCGGCGCGCTCCCGGCGCGCGCT 481
D 1624 AATAACGGAACAAACAAACCTTAACCTTACCAACCGCGCTCCCGAACCGCG 1555
Y 482 GCACCTTCCCGCGGCGCTGGGCGCGGCGGCGGCGCTTCGTCGCGCGCGCGCT 541
D 1564 ACACCTTCCCGCGGCGCTGGGCGCGGCGGCGGCGGCGCTTCGTCGCGCG 1505
Y 542 CTGACCGGAGGAGCGACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
D 1504 CGTAACCGCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1445
Y 602 GTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
```

```
Db 1444 AATAAATAAATCCCTTAACGCTACCAAAAAAATAACGAAATAAATAAATAA 1385
Y 662 AGGTTAATAATCGCCCGCTCCCTTCGCTGCTCTTCATCGAGTTCGCGGAGGCT 721
D 1384 AATATATAATCGCCCGCTCCCTTCGCTGCTCTTCATCGAAATCCGCGAAATCG 1325
Y 722 CGCGCAGCGGAGACACTCTCTCTCGCTCTCCCGGCGAGCGGCGGCTCGGAGCG 781
D 1324 CGCGCAAAACGACACTCTCTCTCGACTCTCCCGACAGGAGGAGGAGGAGGAG 1265
Y 782 TCCGCGGCTCGGGTGCAGCGGCGGCGGCGGCGGCGGCGGAGGATTCGCGGGA 841
D 1264 TCCGAAACTCGAATACAAACGACGAAACGCTTAACGAGAAATTTACCCGGA 1205
Y 842 GTTGTCTCTCTGCTGAGCGGCGGAGCGGCGGCTAGGCGCGGCGGCGGCGGCG 901
D 1204 ATATATCTCTAATCAAAACGCGGAAACGAGCTCAAAACGCGGAAACCGGAG 1145
Y 902 CGAGAGGAGGAGCTTGGCGGCGGCTCTTTGGCGCG 939
D 1144 CGAATAAAGCACTCTACGACGCAATGTTAACCGG 1107
RESULT 11
AX251869/c
LOCUS AX251869 6316 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 130 from Patent WO0168911.
ACCESSION AX251869
VERSION AX251869.1 GI:15985224
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 6316)
AUTHORS Olek.A., Piepenbrock.C. and Berlin.K.
TITLE Diagnosis of diseases associated with the cell cycle
JOURNAL Patent: WO 0168911-A 130 20-SEP-2001;
        Epigenomics AG (DE)
FEATURES
    source
        1..6316
            Location/Qualifiers
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1474 a 255 c 1549 g 3038 t
ORIGIN
Query Match 54.3%; Score 510.8; DB 6; Length 6316;
Best Local Similarity 71.5%; Pred. No. 6.8e-74;
Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;
Y 2 CGGAAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCGAGCTTCCTAAAGG 61
D 2044 CGGAAAAGACGACGACGCTCCCTAAACCTTAACCTTAATTCGCAATCTTC 1985
Y 62 TCCGAAAGACGCTGAGTTCCTCCCTGAGGCTTCCAGTTCAGTTCCTGCTGGCGG 121
D 1984 TACCAACAAACGCTCAATTCCTCCCTCAACGCTTCAATTAATTAACCTTAAC 1925
Y 122 TCCGCTGCTTCCTAGACTTCTCGGACAGCTCTGAAGGGTCTAGGAGCGCGGAC 181
D 1924 TCCGATACCTTCTTAACCTTCTCGAACAATCTAATAAATCAAAACGACGACAC 1865
Y 182 CGGGAAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
D 1864 CGAATAAACAACGACGACGCTCCCTAAACGCTTCAATTAATTAACCTTAAC 1805
Y 242 CGGTGGGAGCGCGGCGGAGCTTCCTTGGATCGGACTTTCGCGCTTAGGCGGCG 301
D 1804 CCGTCGGAAAACGCGACCACTTCCTTAATTCGAACCTTCCGCGCTTAAACCA 1745
Y 302 GCGGAGCTTCAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 361
D 1744 ACGAACTTCAACCTTATCCCTTCCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1685
```

Qy 362 GTGGAGGAGTCTGCAAGGATTTCTGAGCGGATGGGAGGAGGCGGCAAGGCGAA 421
 Db 1684 AATAAAAAAATCTACAAAAATTTCTAAACCGATAAACAAAAAATAAAAAAATA 1625
 Qy 422 GAGGCGGGAGCAAGACCTGACCTGCGGGCGCGCTCCCGGCGCGCGCTCGGCA 481
 Db 1624 AAAAAAGGAAACAAACCTTAACTTACGAAACCGGCTCCCGAACCGGCTCGGCA 1565
 Qy 482 GCACCTTCCCGCGCGCTCGGCGCGCGCGCGCGCGCTCGGCGCGCGCGCGCTCG 541
 Db 1564 ACACCTTCCCGCGCGCTCGGCGCGCGCGCGCGCGCTCGGCGCGCGCGCGCTCG 1505
 Qy 542 GTAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
 Db 1504 CGTAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1445
 Qy 602 GTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 661
 Db 1444 AATAAAAAACCCCTTACGTCACCAAAAAAATAACGAAATAAATAAATAA 1385
 Qy 662 AGGTTAATATGCGCGCGCGCTCGGCTCGCTCTGATCGAGGTCGCGGAGGCTCG 721
 Db 1384 AATAAATAATGCGCGCGCGCTCGGCTCGCTCTGATCGAGGTCGCGGAGGCTCG 1325
 Qy 722 CGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 781
 Db 1324 CGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1285
 Qy 782 TCGGGGCTCGGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 841
 Db 1264 TCGAACTCGAATACAGCAACGAGCGCTAACGAGCAAAATCCCGGAAATA 1205
 Qy 842 GTGTCTCTGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 901
 Db 1204 ATTATCTCTACAAACCGGAAACGAGCGCTCAAAACCGGAAACCGAGCAGCA 1145
 Qy 902 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 939
 Db 1144 CGAAAAACGAACTCTAACGAGCGAAATCGTTAACCGCG 1107

RESULT 12
 AX344257/c 6316 bp DNA linear PAT 01-FEB-2002
 LOCUS
 DEFINITION Sequence 104 from Patent WO020926.
 ACCESSION AX344257
 VERSION AX344257.1 GI:18492145
 WORDS
 IRCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequences.

REFERENCE
 1
 Olek, A., Piepenbrock, C. and Berlin, K.
 TITLE Diagnosis of diseases associated with signal transduction.
 JOURNAL Patent: WO 020926-A 104 03-JAN-2002
 EPIGENOMICS AG (DE)
 FEATURES
 Location/Qualifiers
 source
 1..6316
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1474 a 255 c 1549 g 3038 t

Query Match 54.3%; Score 510.8; DB 6; Length 6316;
 Best Local Similarity 71.5%; Pred. No. 5.8e-74;
 Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

Qy 2 CCGAAAAACAGCAGCAGCGCTCCCTGGGAGCTGAGCTGCTGCGAGCTTCCCAAAAG 61
 Db 2044 CCGAAAAACAGCAGCAGCGCTCCCTGGGAGCTGAGCTGCTGCGAGCTTCCCAAAAG 1985

Qy 62 TCCAAAGCAAGCGTACGTTCCCTCAGGCGTCCAGGTTTCAGTGCCTTGTGCGAGGGTC 121
 Db 1984 TACCAAAACAAACGTCATTCCTCAGGCGTCCAGGTTTCAGTGCCTTGTGCGAGGGTC 1925
 Qy 122 TCCGTCGCTTCTAGACTTCTCGGAGCAGTCTGAAGGTTGAGGAGCGCGGGAGCGG 181
 Db 1924 TCGATACCTTCTTAACTTCTCGAAACAACTTAAANAATCAAAACGACGAAACAACG 1865
 Qy 182 CGGGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 241
 Db 1864 CGAAAAAACAACAAAAAACAACGAACTACGCTCAATCTCCCTCGTACCAAAACA 1805
 Qy 242 CCGTCGCGAGGCGCGCGCGAGCTTCCCTTGGATCGGACTTTTCGCGCCCTTGGCGGCG 301
 Db 1904 CCGTCGCGAAACCGGACCACTTCCCTTAAATCGAACTTTCCGCGCCCTTAAACCAACG 1745
 Qy 302 GCGGAGCTTTCAGGCTTCTCCCTTCCCGAGTTTCGGGCGCGCGCGCGCGCGCGG 361
 Db 1744 ACGAATCTTCACTTATCCCTTCCCGATTTCCGACGACCCCAAACTAATAACCG 1585
 Qy 362 GTGGAGGAGTCTGCAAGGATTTCTGAGCGCGATGGGAGGAGGAGGAGGAGGAGG 421
 Db 1684 AATAAAAAAATCTACAAAAATTTCTTAAACGCGATAAACAAAAAACAACAAAAA 1625
 Qy 422 GAGGCGCGGAGCAAGACCTTGAACCTCGCGGCGCGCTCCCGGCGCGCGCGCGCG 481
 Db 1624 AAAAAACGAAACAAACCTTAACTTACGAAACCGGCTCCCGAACCCCGCTCGGCA 1565
 Qy 482 GCACCTTCCCGAGCGCTCGGCGCGCGCGCGCGCGCTCGGCGCGCGCGCGCGCGCT 541
 Db 1564 ACACCTTCCCGAGCGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1505
 Qy 542 CGTAGCGCGAGGAGCGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
 Db 1504 CGTAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1445
 Qy 602 GTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 661
 Db 1444 AATAAAAAACCCCTTACGTCACCAAAAAAATAACGAAATAAATAAATAA 1385
 Qy 662 AGGTTAATATGCGCGCGCGCTCGGCTCGCTCTGATCGAGGTCGCGGAGGCTCG 721
 Db 1384 AATAAATAATGCGCGCGCGCTCGGCTCGCTCTGATCGAGGTCGCGGAGGCTCG 1325
 Qy 722 CGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 781
 Db 1324 CGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1265
 Qy 782 TCGGGGCTCGGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 841
 Db 1264 TCGAACTCGAATACAGCAACGAGCGCTAACGAGCAAAATCCCGGAAATA 1205
 Qy 842 GTGTCTCTGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 901
 Db 1204 ATTATCTCTACAAACCGGAAACGAGCGCTCAAAACCGGAAACCGAGCAGCA 1145
 Qy 902 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 939
 Db 1144 CGAAAAACGAACTCTAACGAGCGAAATCGTTAACCGCG 1107

RESULT 13
 AX348648/c 6316 bp DNA linear PAT 06-FEB-2002
 LOCUS
 DEFINITION Sequence 106 from Patent WO020807.
 ACCESSION AX348648
 VERSION AX348648.1 GI:18614683
 WORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequences.
 REFERENCE 1
 AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.

Diagnosis of diseases associated with cell signalling
Patent: WO 020907-A 106 10-JAN-2002;
Epigenomics AG (DE)
FEATURES
Source
1. 6316
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 1474 a 255 c 1549 g 3038 t
ORIGIN
Query Match 54.3%; Score 510.8; DB 6; Length 6316;
Best Local Similarity 71.5%; Pred. No. 6.8e-74;
Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;
QY 2 CCGAAAAGACAGCAGCGTCCCTGGCACTGAGTGGTTCAGTCTTCCCAAGG 61
DB 2044 CCGAAAAGACAGCAGCGTCCCTGGCACTGAGTGGTTCAGTCTTCCCAAGG 1985
62 TCGCAGCAGCGTCTAGTCTCCCTAGGCGCTCCAGTTCAGTCTTGGCGAGGTC 121
1984 TACCAACAGCGTCAATTCCTCCCTCAACGCTCCAAATTCATACCTTATACCGAAATC 1925
QY 122 TCGGTGCTTCTAGACTCTCGGACAGTCTGAAGGGTTCAGAGCGGGGACAGG 181
DB 1924 TCGGTGCTTCTAGACTCTCGGACAGTCTGAAGGGTTCAGAGCGGGGACAGG 1863
QY 182 CGGAGAGCAGCAGGAGGAGCAGCGGAGTTCGCTCAGTCTCTCGTTCGCAAGACA 241
DB 1864 CGGAGAGCAGCAGGAGGAGCAGCGGAGTTCGCTCAGTCTCTCGTTCGCAAGACA 1805
QY 242 CCGTCGGAGCGCGGCGCAGCTTCCCTTGGATCGGACTTTCGGCTTAGGCGGCG 301
DB 1804 CCGTCGGAGCGCGGCGCAGCTTCCCTTGGATCGGACTTTCGGCTTAGGCGGCG 1745
QY 302 CGCGAGCTTCAGCTTTCCTTCCCTGAGTTCGGGCGCGCCAGAGCTGAGTAAGCG 361
DB 1744 CGCGAGCTTCAGCTTTCCTTCCCTGAGTTCGGGCGCGCCAGAGCTGAGTAAGCG 1585
QY 362 GTGGGAGGAGTCTGAGGAGTTCCTGAGCGGATGGCAGGAGGAGGCGAGGCGAA 421
DB 1684 AATAAAAGATCTACAAAATTTCTTAAACGGGATTAACAAAAGGAGGAGGAGG 1525
QY 422 GAGGCGCGGAGCAGCAGCTGAGCTGCGGCGCGCGCTTCCGGCGCGGCTCGCA 481
DB 1624 AAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1565
QY 482 GAGCTTCCGCGCGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 541
DB 1564 ACAGCTCCGCGCGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 1505
QY 542 CGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
DB 1504 CGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1445
QY 602 GTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 661
DB 1444 AATAAAAGATCTACAAAATTTCTTAAACGGGATTAACAAAAGGAGGAGGAG 1385
QY 662 AGGTATTAATCGCGCGCGCTTGGCTGCTTCTTCTAGCTTCCGCGGAGGCTCGAG 721
DB 1384 AATATAATATCGCGCGCGCTTGGCTGCTTCTTCTAGCTTCCGCGGAGGAGGAG 1325
QY 722 CGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 781
DB 1324 CGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1265
QY 782 TCGGCGGCTCGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 841
DB 1264 TCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1205
QY 842 GTGTCTCTGCTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 901
DB 1509 CGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1450

DB 1204 ATATCTCTTACTTAACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1145
QY 302 CGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 939
DB 1144 CGAAAAGACGAGCTCTACGACCGAATCGTTAACCGG 1107
RESULT 14
AX458548/c
LOCUS
DEFINITION
Sequence 94 from Patent WO0246454.
ACCESSION
AX458548
VERSION
AX458548.1 GI:21725212
KEYWORDS
SOUP-2
synthetic construct,
synthetic construct
artificial sequences.
CF-ANISM
REFERENCE
1. 6321
AUTHORS
Schacht, O.
TITLE
Diagnosis of diseases associated with angiogenesis
JOURNAL
Patent: WO 0246454-A 94 13-JUN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
Source
1. 6321
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1474 a 258 c 1553 g 3036 t
ORIGIN
Query Match 54.3%; Score 510.8; DB 6; Length 6321;
Best Local Similarity 71.5%; Pred. No. 6.8e-74;
Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;
QY 2 CCGAAAAGACAGCAGCGTCCCTGGCACTGAGTGGTTCAGTCTTCCCAAGG 61
DB 2049 CCGAAAAGACAGCAGCGTCCCTGGCACTGAGTGGTTCAGTCTTCCCAAGG 1990
QY 62 TCGCAGCAGCGTCTAGTCTCCCTAGGCGCTCCAGTTCAGTCTTGGCGAGGTC 121
DB 1985 TACCAACAGCGTCAATTCCTCCCTCAACGCTCCAAATTCATACCTTATACCGAAATC 1930
QY 122 TCGGTGCTTCTAGACTCTCGGACAGTCTGAAGGGTTCAGAGCGGGGACAGG 181
DB 1929 TCGGTGCTTCTAGACTCTCGGACAGTCTGAAGGGTTCAGAGCGGGGACAGG 1870
QY 182 CGGAGAGCAGCAGGAGGAGCAGCGGAGTTCGCTCAGTCTCTCGTTCGCAAGACA 241
DB 1869 CGAAAAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1810
QY 242 CCGTCGGAGGCGGCGGCGAGCTTCCCTTGGATCGGACTTTCGGCGCGGCTCGCA 301
DB 1909 CCGTCGGAGGCGGCGGCGAGCTTCCCTTAAATCGAATTCGCGCGCTTAACGAGG 1750
QY 302 CGGAGCTTCTAGCTTGTCCCTTCCCAATTCGCGGCGCGCGCGCGCGCGCGGAG 361
DB 1749 ACGAAGCTTCTAGCTTATCCCTTCCCAATTCGAGGAGGAGGAGGAGGAGGAGGAG 1690
QY 352 GTGCGGAGGAGTCTGCAAGGATTTCTTCAAGCGATGGCAGGAGGAGGAGGAGGAG 421
DB 1699 AATAAAAGATCTACAAAATTTCTTAAAGCGGATTAACAAAAGGAGGAGGAGGAG 1630
QY 422 GAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
DB 1629 AAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1570
QY 482 GAGCTTCCGCGCGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 541
DB 1569 ACAGCTCCGCGCGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 1510
QY 542 CGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
DB 1509 CGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1450

* 25655 27062: contig of 1408 bp in length
* 27063 27162: gap of unknown length
* 27163 28572: contig of 1410 bp in length
* 28573 28572: gap of unknown length
* 28573 29300: contig of 1258 bp in length
* 29301 30030: gap of unknown length
* 30031 32527: contig of 2497 bp in length
* 32528 32527: gap of unknown length
* 32528 33881: contig of 1254 bp in length
* 33882 33881: gap of unknown length
* 33882 33982: contig of 2632 bp in length
* 33982 36713: gap of unknown length
* 36714 38530: contig of 1817 bp in length
* 38531 38530: gap of unknown length
* 38531 40221: contig of 1591 bp in length
* 40222 40221: gap of unknown length
* 40222 42216: contig of 1895 bp in length
* 42217 42216: gap of unknown length
* 42217 44902: contig of 2586 bp in length
* 44903 45002: gap of unknown length
* 45003 47520: contig of 2518 bp in length
* 47521 47520: gap of unknown length
* 47521 49384: contig of 1754 bp in length
* 49385 49484: gap of unknown length
* 49485 51442: contig of 1958 bp in length
* 51443 51542: gap of unknown length
* 51543 54024: contig of 2482 bp in length
* 54025 54124: gap of unknown length
* 54125 56701: contig of 2577 bp in length
* 56702 56801: gap of unknown length
* 56802 58164: contig of 1363 bp in length
* 58165 58264: gap of unknown length
* 58265 60579: contig of 2415 bp in length
* 60580 60779: gap of unknown length
* 60780 63488: contig of 2709 bp in length
* 63489 63588: gap of unknown length
* 63589 65928: contig of 2340 bp in length
* 65929 66028: gap of unknown length
* 66029 69193: contig of 3165 bp in length
* 69194 69293: gap of unknown length
* 69294 70340: contig of 1047 bp in length
* 70341 70440: gap of unknown length
* 70441 72200: contig of 1760 bp in length
* 72201 72301: gap of unknown length
* 72301 76378: contig of 4078 bp in length
* 76379 76478: gap of unknown length
* 76479 80199: contig of 3721 bp in length
* 80200 80299: gap of unknown length
* 80300 82182: contig of 1883 bp in length
* 82183 82282: gap of unknown length
* 82283 84734: contig of 2452 bp in length
* 84735 87943: gap of unknown length
* 87944 88043: gap of unknown length
* 88044 90121: contig of 2078 bp in length
* 90122 90221: gap of unknown length
* 90222 92597: contig of 2376 bp in length
* 92598 92597: gap of unknown length
* 92598 97180: contig of 4483 bp in length
* 97181 97380: gap of unknown length
* 97381 99751: contig of 2471 bp in length
* 99752 99851: gap of unknown length
* 99852 104895: contig of 5044 bp in length
* 104896 104995: gap of unknown length
* 104996 108453: contig of 3468 bp in length
* 108454 108563: gap of unknown length
* 108564 111925: contig of 3362 bp in length
* 111926 112025: gap of unknown length
* 112026 115621: contig of 3596 bp in length
* 115622 115721: gap of unknown length
* 115722 118437: contig of 2716 bp in length
* 118438 118537: gap of unknown length
* 118538 123236: contig of 4599 bp in length

FEATURES

Query Match

40.1%; Score 377; DB 2; Length 144090;

Best Local Similarity 74.0%; Pred. No. 1.6e-52;

Matches 702; Conservative 0; Mismatches 180; Indels 67; Gaps 15;

Oy 1 SCCGAAAACACGACACGCTCCCTCGGACCTAGCTGTGTTCCAGTCTTCCCAAAG 60
Db 50478 SCCGAAAACACACACGCTCCCTCGGACCTAGCTGTGTTCCAGTCTTCCCAAAG 60419
Oy 51 GTGCCAAGCAA-CGGTCAGTCCCTCAGCGCTCCAGGTTTC-AGTGCCTTGTCCCGAG 118
Db 50418 GTGCTGGCAAGCGCTCCGTTTCCTTCGGAGCTCCGGGTCCAAAGTCCCTTAAGCGGAG 50359
Oy 119 GTCTCCGTCCTTC-TAGACTTCTCGGACACTCTGAAGGGGTCA-GGAGCGCGGGA 176
Db 50358 GTCTCGGTCCTTCCTTCGCGCTCTGGGACACTCTGGCGGGGTCAAGGACCGCGGA 60239
Oy 177 CAGCGCGGAGAGAGGCAAGGAGACACCGGACTCGCGCTCAGTCTTCCGTGCCAA 236
Db 50298 CCGCTC-----AGGAGAGGCTCGACTCGCGCTGTTCTCTCGGTGCCAG 60255
Oy 237 GACACCGTCGCGAGCGCGGCGGCTTCCTTCGAGTGGACTTTCGCGCCCTTAGGGCC 296
Db 60254 GACACCGTCGCGAGCGCGGCGGCGGCTTCCTTCGAGTGGACTTTCGCGCCCTTAGGGCC 50195
Oy 297 AGCGCGCGGAGCTTCAGCTTGTCCCTTCCCGAGTTCGCGCGGCGGCGGCTAGT 356
Db 60194 GCGCGGTGCATTC-----CGCGGTCCTTCCTTCGAGTGGACTTTCGCGCGGCGGCGGCTAGT 50138
Oy 357 AGCGCGGTGCGAGGAG-TCGCAAGGATTCCTGAGCGGATGCGGCGGCGGCGGCGGCA 415
Db 60137 AGCGCGGTGCGAGGAGCTGGCAAGGATTCCTGAGCGGATGCGGCGGCGGCGGCGGCGG 60081
Oy 415 GGGCAAGAGCGCGGCGGCAAGAGCTGAACTTCGCGGCGGCGGCGGCTCCCGGCGGCGG 475
Db 60080 GGGCAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60024
Oy 475 TCGCGCAGCTCTCCCGCGCGGCTC--GGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 533
Db 60023 GCGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 59964
Oy 534 CCGCTCTCCGTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 593
Db 59963 CCGCTC-----GGGAGAAAGAGGCTAGGTGGGAGCG 59928
Oy 594 GATGAGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 653
Db 59927 GATGAGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 59868
Oy 654 CTGGGAAAGGTTTAAATCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 713
Db 59857 CTGGGAAAGGTTTAAATCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 59809
Oy 714 GCTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 771
Db 59808 GCTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 59750
Oy 772 CGGAGCGGCGGCTCCCGGCGGCTCGGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 831
Db 59749 CGGAGCGGCGGCTCCCGGCGGCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 59690
Oy 832 CGGGAAGCTGTTGCTCTCTCGGCTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 891
Db 59689 TCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 59637

Search completed: December 8, 2002, 11:46:32
Job time : 2551.64 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 09:21:42 ; Search time 1219.71 Seconds
(without alignments)
12494.717 Million cell updates/sec

Title: D64016_COPY_500_1440

Perfect score: 941

Sequence: 1 GCCGAAAGACACGACAC.....GCCGGCTTTGGCCCGGG 941

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154056 seqs, 8097743376 residues

Cal number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estimu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_Other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243.4	25.9	697	9	AU136034
2	232.4	24.7	775	13	BI488969 603021595
3	230.8	24.5	906	13	BI820974 603035625
4	230.8	24.5	1171	11	BC029849
5	224.4	23.8	826	9	AL541018 Homo sapi
6	223.4	23.7	972	9	AL543052 AL543052

ALIGNMENTS

RESULT 1
AU136034
LOCUS AU136034 PLACE1 Homo sapiens cDNA clone PLACE1003403 5', mRNA
DEFINITION AU136034
ACCESSION AU136034.1 GI:10996573
VERSION AU136034.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y., Yamamoto.O.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and Isogai.T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project: 5' & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

7	221.8	23.6	1002	9	AL547815
8	215.4	22.9	662	9	AU136704
9	212.2	22.6	1025	9	AL551816
10	211.4	22.5	638	9	AL548886
11	207.4	22.0	860	9	AL546515
12	206.6	22.0	584	14	BQ271447
13	206.4	21.9	1030	9	AL543703
14	204.4	21.7	740	9	AL549158
15	202	21.5	867	9	AL540382
16	200.4	21.3	755	9	AL548841
17	200.4	21.3	1066	13	BI488800
18	197.8	21.0	1067	13	BM546253
19	196.4	20.9	943	13	BI488796
20	173	18.4	828	9	AL546867
21	158.8	16.9	468	9	AL544285
22	103.4	11.0	989	9	AL544285
23	93.8	10.0	615	14	BM735444
24	89.4	9.5	1160	17	AG043473
25	85.4	9.1	925	17	CNS0091P
26	85.2	9.1	925	17	CNS0091P
27	84	8.9	909	9	AL572624
28	79	8.4	811	13	BI888042
29	78.2	8.3	811	13	BI956687
30	78.2	8.3	1288	14	BQ678719
31	77.4	8.2	1137	12	BG809379
32	77	8.2	949	17	AG043499
33	77	8.2	1052	14	BQ950956
34	76.8	8.2	935	17	CNS006XX
35	76.6	8.1	1165	17	AG030649
36	76.2	8.1	1538	17	AG030607
37	75	8.0	1101	17	AG039543
38	74.8	7.9	1151	14	BQ951254
39	74.4	7.9	810	17	AG126142
40	74.4	7.9	1009	17	CNS010EW
41	74.2	7.9	468	14	BQ608819
42	74.2	7.9	948	17	AG161418
43	74.2	7.9	1143	10	AW731158
44	74	7.9	1128	14	BQ10735
45	73.8	7.8	932	17	CNS0072Q

FEATURES
Source

Location/Qualifiers
1. 597
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1003403"
/clone_lib="PLACE1"
/tissue_type="Placenta"
/note="Vector: pME185FL3"
159 a 185 c 220 g 126 t 7 others

BASE COUNT 159 a 185 c 220 g 126 t 7 others
ORIGIN

Query Match 25.9%; Score 243.4; DB 9; Length 597;
Best Local Similarity 99.6%; Pred. No. 4.7e-37;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 597 ATCGAGGTCCCGGGAGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 756

Db 1 ATCGAGGTCCCGGGAGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 60

QY 757 GCAGCGCGCGGCTCGGAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 815

61 GCAGCGCGCGGCTCGGAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 120

QY 817 GCAGCGCGGCTCGGAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 876

Db 121 GCAGCGCGGCTCGGAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 180

QY 877 AGCGCGCGCGGCTCGGAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 936

Db 181 AGCGCGCGCGGCTCGGAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCG 240

QY 937 GCAGCG 941

Db 241 GCAGCG 245

RESULT 2
BI488969
LOCUS 503021695F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5192278 5',
DEFINITION mRNA sequence.
ACCESSION BI488969
VERSION BI488969.1 GI:15328197
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OTHERS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC Http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: ccaps-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1480 row: n column: 23
High quality sequence stop: 775.

FEATURES
SOURCE

Location/Qualifiers
1. 775
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5192278"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPOR16; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT

primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

BASE COUNT 104 a 243 c 324 g 104 t
ORIGIN

Query Match 24.7%; Score 232.4; DB 13; Length 775;
Best Local Similarity 99.6%; Pred. No. 5.9e-35;
Matches 233; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 708 CGGAGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGCGG 767

Db 1 CGGAGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGCGG 60

QY 768 GGCTCGGAGCGGCCAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGG 827

Db 51 GGCTCGGAGCGGCCAGCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGG 120

QY 828 TACCGGGGAGTGGTTGTCTCTCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGG 887

Db 121 TACCGGGGAGTGGTTGTCTCTCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGG 180

QY 888 CGGCGCGCGGCGGAGCGGACACTCTCTCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGG 941

Db 181 CGGCGCGCGGCGGAGCGGACACTCTCTCGGCTCGGAGCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGG 234

RESULT 3
BI820974
LOCUS 603035625F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176591 5',
DEFINITION mRNA sequence.
ACCESSION BI820974
VERSION BI820974.1 GI:159332524
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OTHERS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC Http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: ccaps-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1440 row: a column: 08
High quality sequence start: 2
High quality sequence stop: 765.

FEATURES
source

Location/Qualifiers
1. 906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176591"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPOR16; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb. insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C.

Gruber (Invitrogen). Research Genetics tracking code

621. Note: this is a NIH_MGC Library.

BASE COUNT 219 a 228 c 264 g 194 t

ORIGIN

Query Match 24.5%; Score 230.8; DB 13; Length 906;
Best Local Similarity 99.1%; Pred. No. 1.2e-34;
Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 708 CGGAGGCTCGGAGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGAGCGCGC 767

Db 1 CGGAGGCTCGGAGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGAGCGCGC 50

QY 768 GGTCTGAGCGGGCTCCGGGCTCGGTGAGCGGCCAGCGGCGCTTGGCGGCGAGGAT 827

Db 61 GGTCTGAGCGGGCTCCGGGCTCGGTGAGCGGCCAGCGGCGCTTGGCGGCGAGGAT 120

QY 828 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGCGGAGCGGCTCTAGCGCGCGGG 887

Db 121 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGCGGAGCGGCTCTAGCGCGCGGG 180

QY 888 CCGCGCGCGCGAAGCGAGGAGCGGACTCTGGCGCGCGGGTCTTGGCGCGCGGG 941

Db 181 CCGCGCGCGCGAAGCGAGGAGCGGACTCTGGCGCGCGGGTCTTGGCGCGCGGG 234

RESULT 4

BC029849 1171 bp mRNA linear HTC 20-MAY-2002

LOCUS BC029849 Homo sapiens, clone IMAGE:5176591, mRNA.

DEFINITION BC029849.1 GI:20987442

ACCESSION HTCC

VERSION Homo sapiens.

KEYWORDS Homo sapiens

SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1171)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.

USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgpbbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HQSC

Web site: <http://www.hqsc.bcm.tmc.edu/cdna/>

Contact: ang@bcm.tmc.edu

Gunnaratne, P.H., Garcia, A.M., Lu, X., Hui, K., S.W., Hale, S.M.,

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 51 Row: m Column: 23

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4503748

This clone has the following problem: incomplete processing.

FEATURES Location/Qualifiers

1..1171

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5176591"

/tissue_type="Brain, Lung, Testis, adult, pooled whole"

/clone_lib="NIH_MGC_115"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

BASE COUNT 340 a 260 c 302 g 269 t

ORIGIN

Query Match 24.5%; Score 230.8; DB 11; Length 1171;

Best Local Similarity 99.1%; Pred. No. 1.2e-34;

Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 708 CGGAGGCTCGGAGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGAGCGCGC 767

Db 1 CGGAGGCTCGGAGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGCGAGCGCGC 60

QY 768 GGTCTGAGCGGGCTCCGGGCTCGGTGAGCGGCCAGCGGCGCTTGGCGGCGAGGAT 827

Db 61 GGTCTGAGCGGGCTCCGGGCTCGGTGAGCGGCCAGCGGCGCTTGGCGGCGAGGAT 120

QY 828 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGCGGAGCGGCTCTAGCGCGCGGG 887

Db 121 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGCGGAGCGGCTCTAGCGCGCGGG 180

QY 888 CCGCGCGCGCGAAGCGAGGAGCGGACTCTGGCGCGCGGGTCTTGGCGCGCGGG 941

Db 181 CCGCGCGCGCGAAGCGAGGAGCGGACTCTGGCGCGCGGGTCTTGGCGCGCGGG 234

RESULT 5

AL541018

LOCUS AL541018 LTI_FLO02_P11

DEFINITION AL541018 Homo sapiens cDNA clone CS0DR005Y102 5 prime

ACCESSION AL541018

VERSION AL541018.1

KEYWORDS EST.

SOURCE Human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 826)

AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91005 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES Location/Qualifiers

1..826

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DE005Y102"

/clone_lib="LTI_FLO02_P11"

/lab_host="DH10B"

/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand

cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the

pCMVSPORT 6 vector. Library was constructed by Life

Technologies. Contact : Feng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville,

Maryland 20850, USA Fax : (1) 301 610 8371 Email :

liang@lifetech.com URL :

<http://fulllength.invitrogen.com>

BASE COUNT 201 a 218 c 243 g 162 t

ORIGIN

Query Match 23.8%; Score 224.4; DB 9; Length 826;

Best Local Similarity 99.6%; Pred. No. 2e-33;

Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 716 TCGGAGCGCGCGAGCGGACACTCTCTCGGCTCTCTCCCGGCGAGCGCGCGCTCGGA 775

Db 1 TCGGAGCGCGCGAGCGGACACTCTCTCGGCTCTCTCCCGGCGAGCGCGCGCTCGGA 50

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 09:24:02 : Search time 83.5704 Seconds
(without alignments)
7710.006 Million cell updates/sec

Title: x51602_copy_1900_4000
Perfect score: 2101
Sequence: 1 TATAACACAGTGGCCAA.....CGACAGCAGCAGTGGTGG 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PC105_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	713.8	34.0	4071	US-09-098-707A-1	Sequence 1, Appli
2	713.8	34.0	4071	US-09-483-539-1	Sequence 1, Appli
3	710.6	33.8	4236	US-07-810-116-7	Sequence 7, Appli
4	710.6	33.8	4236	US-07-930-548A-7	Sequence 7, Appli
5	710	33.8	5470	US-08-443-861-1	Sequence 1, Appli
6	710	33.8	5470	US-08-193-829B-1	Sequence 1, Appli
7	708.4	33.7	5406	US-07-813-593-3	Sequence 1, Appli
8	708.4	33.7	5406	US-07-977-451-5	Sequence 3, Appli
9	708.4	33.7	5406	US-07-946-507-3	Sequence 5, Appli
10	708.4	33.7	5406	US-08-252-517-5	Sequence 3, Appli
11	708.4	33.7	5406	US-07-906-397A-5	Sequence 5, Appli
12	708.4	33.7	5406	US-08-601-891-5	Sequence 5, Appli
13	708.4	33.7	5406	US-09-021-324-5	Sequence 5, Appli
14	708.4	33.7	5406	US-09-021-324-5	Sequence 5, Appli
15	708.4	33.7	5406	US-09-021-324-5	Sequence 5, Appli
16	708.4	33.7	5406	US-09-021-324-5	Sequence 5, Appli
17	690	32.8	2352	US-08-232-538-17	Sequence 17, Appli
18	690	32.8	2352	US-08-786-164-17	Sequence 17, Appli
19	624	29.7	2523	US-09-051-363-1	Sequence 1, Appli
20	541.6	25.8	4195	US-08-340-011-1	Sequence 1, Appli
21	541.6	25.8	4195	US-08-901-710-1	Sequence 1, Appli
22	541.6	25.8	4416	US-08-795-730-1	Sequence 1, Appli
23	541.6	25.8	4416	US-09-355-700-1	Sequence 1, Appli
24	541.6	25.8	4416	US-08-601-132-36	Sequence 36, Appli
25	541.6	25.8	4425	US-08-222-616-31	Sequence 31, Appli
26	541.6	25.8	4425	US-08-446-648-31	Sequence 31, Appli
27	541.6	25.8	4425	US-09-04228-31	Sequence 31, Appli

28	541.6	25.8	4795	1	US-08-340-011-3	Sequence 3, Appli
29	541.6	25.8	4795	3	US-08-901-710-3	Sequence 3, Appli
30	541.6	25.8	9108	4	US-08-446-648-45	Sequence 45, Appli
31	541.6	25.8	9108	5	US-08-446-648-45	Sequence 45, Appli
32	321.2	15.3	2313	1	US-08-232-538-5	Sequence 5, Appli
33	321.2	15.3	2313	4	US-09-427-353-1	Sequence 1, Appli
34	321.2	15.3	2651	2	US-08-786-164-5	Sequence 5, Appli
35	306	14.6	551	3	US-08-750-141A-5	Sequence 5, Appli
36	282.8	13.5	6827	1	US-08-222-616-17	Sequence 17, Appli
37	282.8	13.5	6827	4	US-08-446-648-17	Sequence 17, Appli
38	282.8	13.5	6827	5	US-08-446-648-17	Sequence 17, Appli
39	282.8	13.5	6827	5	US-08-446-648-17	Sequence 17, Appli
40	259	12.3	5084	1	US-08-306-691B-21	Sequence 21, Appli
41	249.4	11.9	3475	1	US-08-222-299-3	Sequence 25, Appli
42	249.4	11.9	3475	2	US-08-434-878-3	Sequence 3, Appli
43	249.4	11.9	3475	5	US-08-434-878-3	Sequence 3, Appli
44	249.4	11.9	3476	1	US-08-183-211-1	Sequence 1, Appli
45	249.4	11.9	3476	5	US-08-183-211-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-098-707A-1
: Sequence 1, Application US/09098707A
: Patent No. 5204011
: GENERAL INFORMATION:
: APPLICANT: Kendall, Richard L.
: Thomas, Kenneth A.
: Mao, Xianzhi
: Jebben, Andrew J.
: TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, NDR
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: P.O. Box 2000
: CITY: Rahway
: STATE: NJ
: COUNTRY: US
: ZIP: 07065-0907
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/098,707A
: FILING DATE: 17-Jun-1998
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Hand, J. Mark
: REGISTRATION NUMBER: 36,545
: REFERENCE/DOCKET NUMBER: 19963PV
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 732/594-3905
: TELEFAX: 732/594-4720
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4071 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-098-707A-1

Query Match 34.0% Score 713.8; DB 4; Length 4071;
Best Local Similarity 65.5%; Pred. No. 8.5e-212;
Matches 107; Conservative 0; Mismatches 562; Indels 5; Gaps 2;
QY 218 TCAATGATGTTCCCTGCAAGATTCAGGCACCTATGCTGACAGCCAGGATGTATACA 277
||| ||||| ||| ||||| ||| ||||| ||||| ||| |||

Db 1886 TTAGAATGATCTTGGAGGACCAAGAGACTATGTCCTTCTCAGACAGAGAGA 1845
QY 278 CAGGGAGAAATCTCCAGAGAAGAAATTAATACATCAGATCAGGAAGCACCATACC 337
Db 1946 CCAAGAAAAGACATTCGTGGTTCAGGACCTCAGAGCTCCTAGAGGCTGAGCAGCCACGA 2005
QY 338 TCTGCGAAACCTCAGTGATCAGACAGTGGCCATCAGCAGTTCACACCTTACAGCTGIC 397
Db 2006 TCACAGAAACCTGGAGATCAGACAGCAAGTATGGGAAGCATCGAGTCTCATGCA 2065
QY 398 ATGCTAATGGTGTCCCGAGCCTCAGATCAGTCTGGTTTAAACACACCAAAATACAC 457
Db 2066 CGGCACTGGGAATCCCTCCACAGATCATGTGTTTAAAGATAATGAGACCTTGTAG 2125
QY 458 AAGACCTTGGAAATATTTAGACAGCAGCAAGCAGCAGCCTGTTTATGAGAGCTACAG 517
Db 2126 AAGACTCAGGCAATGATTAAGAGATGGGAACCGGAACCTCAGTATCCGAGAGTAGGA 2185
QY 518 AAGAGATGAAGTGTCTATCAGTCCAAAGCCACCAACAGAGAGGCTCTGGAAAGT 577
Db 2186 AGGAGCAGAGCCTTACACTCCAGGATGCACTGCTGCTGTCGCAAGTGG 2245
QY 578 CAGCATCTCAGTCTTCAAGCACTCGGACAGCTTAATCTGAGCTGATCATCTAA 637
Db 2246 AGGCAATTTTCAATAGAGGTGCCAGGAAGACGAACTTGGAAATCATTTCTAG 2305
QY 638 CATGCACTGTGGCTGGGACTCTTCTGCTCTCTTAATCCCTCTTATCCGAAA 697
Db 2306 TAGGCAGGGGTATGGCATGTTCTTCTGGCTACTTCTGTCATCTTACCTACGAGCC 2365
QY 698 TGAAGAGTCTCTTCTGTAATGAAGTCTACTTACCTATCAATTAATATGAGCACC 754
Db 2366 TTAAGCGGGCAATGAGGGAAGTGAAGACAGAGGTACCTGTCCATCTGTCAGTCA 2425
QY 755 ATGAGATTCCTTGGATGAGCAGTGTAGCGGCTCCCTTAATGATCCAGCAAGTGGAGT 814
Db 2426 ATGAATCTCCATGGATGAACATTTGAGACGACTGCTTAATGATGCCAAATGAGAA 2485
QY 815 TTGCGCGGAGAGACTTAACTGGGCAATCACTTGAAGAGGGCTTTTGAAGATGG 874
Db 2486 TCCGAGAGCGGTGACCTAGTAGAGCTCTTGGCGGTGGCTTTGGCAAGTGA 2545
QY 875 TTAAGCATCAGATTTGGCATTAAGAATCACCCTGTCGCGGAGTGTGGCTGTGAAA 934
Db 2546 TTAAGCAGAGATGCTTTGGAATTAACAAGACAGCAACTTGCAGGACAGTAGCAGTCA 2605
QY 935 TCTGAAGAGGCGCCAGCGGCGAGTACAAAGCTCTGATGAGCTGATTAATCT 994
Db 2606 TGTGAAGAGAGAGACACACAGCAGTGAAGCACTGAGCTCTCTGCTGAGCTCAAGATCC 2665
QY 995 TGACCCCATTTGGCCACCTATCTGAACGTGGTTAACCTCTGGGAGCTTCCACCAAGCAAG 1054
Db 2666 TCATTCATTTGCTACCACTCAATGTGTCAACCTTCTAGGTGGCTTACCAAGCCAG 2725
QY 1055 GAGGCTCTGATGCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1114
Db 2726 GAGGCTCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2785
QY 1115 AGAGCAAACTGATCTTATTTTCTCAACAAGGATGAGCAGCTACACATGAGAGCTAAGA 1174
Db 2786 GAGCAGAGAGAAATGATTTGCTCTACAGACCAAGGGGCAAGATCTGCTGAGGGA 2845
QY 1175 AAGAAAAATGAGCCAGGCTGGAACAGGCAAGCAAAACCAACAGCAGTATAGCTGACCA 1234
Db 2846 AAGACTAGTTGGAGCAATCCCTG---TGGATCTGAAGAGGCTTGGACAGCATCA 2902
QY 1235 GAGGCAAGCTTTGAGGCTCGGCTTTTACGAGAGATTAAGCTGAGTGTGAGT 1294
Db 2903 GTAGCAGAGCTCAGCAGCTGCTGATTTGTTGAGGAGAGTCCCTCAGTGTAGAG 2962
QY 1295 AAGAGGAGTCTGAGCGTTTCTACAGGAGCCATCAGTATGAGAGATGATTTCT 1354
Db 2963 AAGAGGAGCTCTGAGATCTGATAGGACTTCTGACCTTGGAGCACTCATCTGTT 1422

RESULT 2

US-09-483-539-1
; Sequence 1, Application US/09483539
; Patent No. 6359115
; GENERAL INFORMATION:
; APPLICANT: Kendall, Richard L.
; APPLICANT: Thomas, Kenneth A.
; APPLICANT: Mao, Xianzhi
; APPLICANT: Tennen, Andrew J.
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Merck & Co., Inc.
; STREET: P.O. Box 2000
; CITY: Kenilworth
; STATE: NJ
; COUNTRY: US
; ZIP: 07055-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/483,539
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19963PV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 1:

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/810,116

FILING DATE: 25-FEB-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/330,548

FILING DATE: 23-NOV-1992

ATTORNEY/AGENT INFORMATION:

NAME: Gordon, Alan M.

REGISTRATION NUMBER: 30,637

REFERENCE/DOCKET NUMBER: 31,298-01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-831-3244

TELEFAX: 201-831-3305

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 4236 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..4068

US-08-810-116-7

Query Match 33.8%; Score 710.6; DB 1; Length 4236;

Best Local Similarity 65.3%; Pred. No. 8.6e-211;

Matches 1075; Conservative 0; Mismatches 564; Indels 5; Gaps 2;

```
QY 218 TCATGAATGTTCCCTGCAAGATTACGACCTATGCTCAGACGACGAGGATGATATACA 277
DB 1886 TTAAGATGCAATCTTCAGGACCCAGGAGACTATGCTCCCTTGCTCAAGCAGGANGA 1945
QY 278 CAGGGGAGAAATCTCCAGAGAGAAATATACAAATCAGAGATCAGGAGCACCATACC 337
DB 1946 CCAAGAAAGACATTCGGTGGTTCAGGACCTCAGAGCTTAGAGCGTGTGGCACCACGA 2005
QY 338 TCTCCGAAACCTCAGTGATCACAGATGGCCATCAGGATTCACACCTTACACTGTC 397
DB 2006 TCACAGAACTGGAGATCAGACGACAGATATGGGAAAGCATCGAAGTCTCATGCA 2065
QY 398 ATGCTAATGTTGTCCTGAGCTCAGATCCTTGTGTTTAAACAAACCAACAAATACAC 457
DB 2066 CGGCATCTGGAAATCCCTCCACAGATCATGTGTTTAAAGATAATGAGACCTTGTAG 2125
QY 458 AAGACCTGGATTAATTTAGGACCCAGGACGACGCTGTTTATGAAAGATCACAG 517
DB 2126 AAGACTCAGGCAATGATGAGGATGGGACCGGACCTCCTATCCGACAGGTGAGGA 2185
QY 518 AAGAGGATGAAGTGTCTATCACTGCAAGCCCAACCAAGAGGCTCTGTGAAAGTT 577
DB 2186 AGGAGGAGAGGCTTACACCTGCCAGGATGCGAGTTCCTGCTGTCMAAGTGG 2245
QY 578 CAGCAATCTACTGTTCAAGACCTCGGACAGCTTAATCTGAGCTGATCCTTAA 637
DB 2246 AGGCAATTTTCAATAGAGAGTGGCCAGGAAAGACCAATTCGAAATCATTTCTAG 2305
QY 638 CATGACCTGTGGCTGGAGCTCTCTCTGCTCTTAAACCTCTTATCCGMAAA 697
DB 2306 TAGGCACAGGATGATGCAATTTCTTGGCTACTTCTGCTATCTCTCTAGGACCG 2365
QY 698 TGAAGAGTCTCTTCTTCTGAAATTAAGACTGACATCTATCAATTAATTAAGACCCAG 754
DB 2366 TTAAGCGGCAATGAGGAGGAACTGACAGAGGCTTCTGCTATCTCTCTAGGACCG 2425
QY 755 ATGAGATCTCTTGGATGACAGCTGTAGCGGCTCTCTTATGATGCCAGAGTGGAGT 814
DB 2426 ATGAATCTCTTGGATGAAATTTGTAACGACTGCTTATGATGCCAGAAATGGAAAT 2485
QY 815 TTGCCCGGAGAGACTTAACATCGGCAATCACTTGAAGAGGCTTTTGAAGTGG 874
```

RESULT 4

US-07-930-548A-7

Sequence 7, Application US/07930548A
Patent No. 5861301
GENERAL INFORMATION:
APPLICANT: Terman, Bruce I.
APPLICANT: Carrion, Miguel E.
TITLE OF INVENTION: Identification of a No. 5861301el Human Growth
TITLE OF INVENTION: Factor Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/930.548A
FILING DATE: 23-NOV-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30.637
REFERENCE/DOCKET NUMBER: 31,298-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4236 base pairs
TYPE: nucleic acid
STRAINEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4068
US-07-930-548A-7

Query Match 33.8%; Score 710.5; DB 2; Length 4236;
Best Local Similarity 65.3%; Pred. No. 8.6e-211;
Matches 1075; Conservative 0; Mismatches 564; Indels 6; Gaps 2;
218 TCATGAATGTTCCCTGCGAAGATTCAGGCACCTATGCTGCGAGCCAGGAATGATACA 277
1886 TTAAGATGATCTTGTGAGGACCAAGGAGATATGCTGCTTGTCTCAAGCAGGAGA 1545
278 CAGGGAGAGAAATCTTCAGAGAGAGAAATTAACATCAGATCAGAGCAGCACCATACC 337
1946 CCAAGAAAGACATTCGGTGGTTCAGGAGCTCAGAGTCTAGAGGCTGTGCGACCCACGA 2005
338 TCTCGGAAACCTTCAGTATCAGAGTGGCCATCAGAGTTCACCACTTACAGCTGTC 397
2006 TCACAGAACTTGGAGATCAGACAGCAAGATATGGGAAACCATCGAGTCTATGCA 2065
398 ATGCTAATGTTGCCCGGCTCAGATCAGTGGTTTAAACAAACACCAAAATACAAAC 457
2066 CGGCACTGGGAATCCCTCCACAGATCATGTGTTTAAAGATATGAGACCCCTTGAG 2125
458 AAGAGCTTGAAGATATTTAGACAGCAGAGCAGCAGCTGTTTATTAAGAGATCAGAG 517
2126 AAGACTCAGGATTTGATTAAGAGATGGGAACCGGAACTCAGTATCCGAGAGTGAGGA 2185
518 AAGAGATGAGGTGCTATCAGTCAAGAGCCACCAACAGAGAGGCTCTGTGGAAAGTT 577
2186 AAGAGAGAGAGGCTCTACAGCTCCAGGATGCGAGTGTCTTGTGCTGTGCAAAAGTG 2245
578 CAGCATACCTCTGTTCAAGGAACCTCGGACAGAGCTATATCTGAGGATGATCACTTAA 637

Db 2246 AGGCAATTTTCAATAGAGGTGCCAGAGAAAGACGAACTTGGAAATCATTTTCTAG 2305
QY 638 CATGACCTGTGTGGCTGCGACTCTCTCTGCTCTTATTAACCTCTCTTATTCGAAAA 697
Db 2306 TAGGACAGCAGCGATGATGCAATGTTCTCTGCTGCTTCTGTGCTATCTCTGCTAGGAC 2365
QY 698 TGAAGAGTIC---TTCCTTGAATAAAGACTGACTACCTATCAATTAATTAATGACCCAG 754
Db 2366 TTAAGCGGCAATGGAGGGAACCTGAAGACAGGCTACTTGTCCATGCTCATGATCCAG 2425
QY 755 ATGAAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCTCTTATGATGCCAGAGTGGAGT 814
Db 2425 ATGAAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCTCTTATGATGCCAGAGTGGAGT 814
QY 815 TTGCGCGGAGAGACTTAACTGGGCAATCACTTGGGAGAGGGCTTTTGGAAAGTGG 874
Db 2486 TCCCGACAGACCGGCTGAACCTAGTGAAGCTCTTGGCGTGTGCTCTTGGCCAGAGA 2545
QY 875 TTAAGCATCAGCATTTGGCATTAAGAAATCACTAGTGCAGGACTGTGCTGTGAAGA 934
Db 2545 TTGAAGCAGATGCTTTGGAATTGACAAGACAGCAACTTGCAGGACAGTAGCAGTCAAA 2605
QY 935 TGTGAAGAGGGGGCCAGCCGACGAGTACAAGCTCTGTGACTGAGCTAATAATCT 994
Db 2606 TGTGAAGAGGGGGCCAGCCGACGAGTACAAGCTCTGTGACTGAGCTAATAATCT 994
QY 995 TGACCCACATTTGGCCACCATCTGAACCTGTTTAACTGCTGGAGAGCTGCACCAAGCAAG 1054
Db 2666 TCATTATATTTGGTCAACATCTCAATGTTGTTGCTCACTTCTAGTGCCTGTACCAAGCAG 2725
QY 1055 GAGGCTCTGTGATGTTGATTTGAATCTGCAATATGGAATCTCTCAACTACCTCA 1114
Db 2726 GAGGCTCTGTGATGTTGATTTGAATCTGCAATATGGAATCTCTCAACTACCTCA 2785
QY 1115 AGAGCAAGCTGACTTATTTTCTCAAGAGTGCAGCTACACATGAGGCTAAGA 1174
Db 2786 GAGCAAGAGAGAGATTTTCTCAAGAGTGCAGCTACACATGAGGCTAAGA 1174
QY 1175 AAGAAAAATGGAGCCGCTGGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1234
Db 2846 AAGACTAGTTGGAGCAATCCCTG---TGATCTGAAGCGGCTTGCACAGCATCACA 2902
QY 1235 CGAGGAAAGCTTTGGAGCTCGGCTTTTCAAGAGATTAAGAGTCTGAGTGTGAGG 1294
Db 2903 GTAGCCAGAGCTCAGGAGCTCTGATTTTGTGAGAGAGTCCCTCAGTGTATAGAAG 2962
QY 1295 AAGAGAGGATTTCTGAGGCTTTTCTAAGAGAGCCATCATATGGAAGATCTGATTTCT 1354
Db 2963 AAGAGAGGATTTCTGAGGCTTTTCTAAGAGAGCCATCATATGGAAGATCTGATTTCT 1354
QY 1355 ACAGTTTCAAGTGGCCAGAGCATGAGTCTCTGCTTCCAGAAAGTGCATTCATCGG 1414
Db 3023 ACAGTTTCAAGTGGCCAGAGCATGAGTCTCTGCTTCCAGAAAGTGCATTCATCGG 3082
QY 1415 ACTTGGAGGAGAGATTTCTTATCTGAGAAACAGTGTGAGATTTGATTTG 1474
Db 3083 ACTTGGAGGAGAGATTTCTTATCTGAGAAACAGTGTGAGATTTGATTTG 3142
QY 1475 GCTTGGCCGCGATTTTATAGAAACCCGATTTATGTGAGAAAGAGATCTCGACTTC 1534
Db 3143 GCTTGGCCGCGATTTTATAGAAACCCGATTTATGTGAGAAAGAGATCTCGACTTC 3202
QY 1535 CTCGAAATGAGTGGCTCCGAGATCTCTTGTGACAAATCTACAGCAGCAGAGGAGC 1594
Db 3203 CTCGAAATGAGTGGCTCCGAGATCTCTTGTGACAAATCTACAGCAGCAGAGGAGC 3262
QY 1595 TGTGCTCTTGGAGATTTGCTGTGGAAATCTCTCTTGTGAGTGGCTCTCCATCCAG 1654
Db 3263 TGTGCTCTTGGAGATTTGCTGTGGAAATCTCTCTTGTGAGTGGCTCTCCATCCAG 3322
QY 1655 GAGTACAAATGAGTGGAGCTTTTTCAGTCCCTGTGAGGAGGAGCATGAGATGAGCTC 1714
Db 3323 GGTAAAGATGATGAAGATTTTGTAGGCGATTTGAAAGAGGAGTGAAGTGAAGGCC 3382

QY 1715 CTGAGTACTTCTTCTGAAATCTATGATGATGCTGGAGCTCTGGACAGAGACCCAA 1774
Db 3383 CTGATTATATACACCAAGATGATACAGACCATCTCTGGACTCTGGACGGGGGCCA 3442
QY 1775 AGAAGGCCAGATTTGAGAGACTTGGAAAGCTAGTGATTTGCTTCAAGCAATG 1834
Db 3443 GTCAGAGACCCAGCTTTTCAAGTGTGGTGAACATTTGGGAATCTTTGCAAGCTAATG 3522

QY 1835 TACAACAGGATGTAAGACTACAT 1859
Db 3503 CTCAGCAGGATGCAAGACTACAT 3527

RESULT 5

US-08-443-861-1

: Sequence 1, Application US/08443851

: Patent No. 5851999

: GENERAL INFORMATION:

: APPLICANT: Cullrich, Axel

: APPLICANT: Bisau, Werner

: APPLICANT: Millauer, Birgit

: APPLICANT: Gazit, Aviv

: APPLICANT: Levitzki, Alex

: TITLE OF INVENTION: Plk-1 Is A Receptor For Vascular

: TITLE OF INVENTION: Endothelial Growth Factor

: NUMBER OF SEQUENCES: 5

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Pennie & Edmonds

: STREET: 1155 Avenue of the Americas

: CITY: New York

: STATE: New York

: COUNTRY: U.S.A.

: ZIP: 10036-2711

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/443,861

: FILING DATE: 22-MAY-1995

: CLASSIFICATION: 514

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/193,829

: FILING DATE: 09-FEB-1994

: ATTORNEY/AGENT INFORMATION:

: NAME: Coruzzi, Laura A.

: REGISTRATION NUMBER: 30,742

: REFERENCE/DOCKET NUMBER: 7683-060

: TELEPHONE: (212)790-9090

: TELEFAX: (212)859-9741

: TELEX: 66141 PENNIE

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 5470 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: unknown

: TOPOLOGY: unknown

: MOLECULE TYPE: DNA

: FEATURE:

: NAME/KEY: CDS

: LOCATION: 286..4386

US-08-443-861-1

Query Match 33.8%; Score 710; DB 2; Length 5470;

Best Local Similarity 55.8%; Pred. No. 1.6e-210;

Matches 1081; Conservative 0; Mismatches 575; Indels 6; Gaps 2;

QY 201 CACTCTTAATCTTACCAATGATTTTCCCTGCAAGATTCAGGCAACCTATGCTGCGAG 2
Db 2148 CACTCTTAATCTTACCAATGATTTTCCCTGCAAGATTCAGGCAACCTATGCTGCGAG 217

QY 1338 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1397
DB 3295 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3344
QY 1398 AAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1457
DB 3345 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3404
QY 1458 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1517
DB 3405 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3464
QY 1518 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1577
DB 3465 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3524
QY 1578 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1637
DB 3525 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3584
QY 1638 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1697
DB 3585 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3644
QY 1698 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1757
DB 3645 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3704
QY 1758 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1817
DB 3705 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3764
QY 1818 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 1877
DB 3765 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATCTGAGTTCTCTGCTCCAG 3836
RESULT 6
US-08-193-8298-1
Sequence 1, Application US/081938298
Patent No. 6177401
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Risa, Werner
APPLICANT: Millauer, Birgit
APPLICANT: Gazit, Aviv
APPLICANT: Levitzki, Alex
TITLE OF INVENTION: FIX-1 Is A Receptor For Vascular
TITLE OF INVENTION: Endothelial Growth Factor
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fennie & Edmonds
STREET: 1155 Avenue Of The Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 09-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-060
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5470 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 286..4395
US-08-193-8298-1

Query Match 33.8%; Score 710; DB 4; Length 5470;
Best Local Similarity 65.0%; Pred. No. 1.6e-210;
Matches 1081; Conservative 0; Mismatches 575; Indels 6; Gaps 2;

QY 201 CACCTTAACTTACCATCATGATGTTTCCCTGCAAGATTCAGGCACCTATGCTGCGAG 250
DB 2148 CATCTTGAATGGCAATTCAGATGGCTCTCTGAGGACCAAGGCGACTATGTTGCTC 2207
QY 261 AGCCAGGATGTATACACAGGGGAAGAAATCTCCAGAGAAAGAAATTAACAATCAGAGA 320
DB 2208 TGCTCAAGATAAGACCAAGAAAGACATTCCTGGTCAACAGCTCATCATCTCTAGA 2267
QY 321 TCAGAGAGCACCATATCTCTGCGAACTCAGTGATCAGACAGTGGCCATCAGCAGTTC 380
DB 2268 GCGATGCGCCCATCATCACCAGGAAATCTGGAGATTCAGACAAACCAATTCGCGAGAC 2327
QY 381 CACCACTTATAGATGTGCTAAATGGTGTCCCGAGGCTTCAGATCAGTCTGTTTAAAA 440
DB 2328 CATTCAGAGTATTCAGGATCTGGAATCTTCCACACATTCATGTTCAAGA 2387
QY 441 CAGCACAATAATACACAGAGCCCTGGAATTTTAGACAGGAGAACACAGCAGCTGTT 500
DB 2388 CAAGGAGACCCCTGGTAAAGATTCAGGCAATTCAGGATTCAGGATTCAGGATTCAG 2447
QY 501 TATTGAAGAGTACAGAGAGGATGAGGTGCTATCTGCAAGCCACCAACAGAA 560
DB 2448 TATCGGAGGTGAGGAGAGGATGAGGCTCTACCTGCGAGGCTCTGCAATGCTCT 2507
QY 561 GGGCTCTGGAAGTTCAGATACCTCAGTCTTCAAGGAGTCTGGAAGTCTTAACT 620
DB 2508 TGGCTGTGCAAGAGGAGGAGGCTCTTCATATAGAGGTGCCAGGAAAGACCACT 2567
QY 521 GAGCTGATCAGTCAATGAGTACCTGTGGCTGCGACTCTCTTGGCTCTTATTAAC 580
DB 2568 GGAAGTCAATTAATCTGCGGAGTGCAGTGCAGTGTCTCTGCTCTCTCTGCTCT 2627
QY 581 CTCTCTTATCCGAAATTAAGAAAGTCTCTCTGTAATTAAGAAAGTCTCTCTCTCT 737
DB 2628 CATTCGCTAGGAGGCTTACGCGGCAATGAGGAGGAGTGAAGACAGAGTCTGCT 2687
QY 738 AATTATAAGGAGGAGGAGTCTTCTTGGATGAGTGTGAGCGGCTCTCTTATGA 797
DB 2688 TATTGATGAGTCCAGATGAATTCCTGATGAGGCTGTGAACGCTTCTGCTTATGA 2747
QY 798 TCCAGGAGTGGGAGTTCCTCCGAGAGAGTCTTAACTGCGCAATCAGTGAAGAG 857
DB 2748 TCCAGGAGTGGGAGTTCCTCCGAGAGGAGTCTTAACTGCGCAATCAGTGAAGAG 2807
QY 858 GGCCTTTGGAAGTGGTTCAGGATTCAGGATTCAGGATTCAGGATTCAGGATTCAGG 917
DB 2808 TGCCTTCGCGGAGGATTCAGGAGGAGGCTTTCAGGATTCAGGAGGAGGAGTCTG 2867
QY 918 GACTGTGGCTGTGAATGCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
DB 2868 AACAGTGGCTGTGAATGCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2927
QY 978 GACTGTGGCTGTGAATGCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1037
DB 2928 GCTGAGTCAAGATCTCTATCAGGATTCAGGATTCAGGATTCAGGATTCAGGATTCAG 2987

QY	1038	AGCTGCACCAAGCAGGAGGCGCTCTGATGGTGTGATTTGAAATCTCAAAATATGGAAA	1097
Db	2988	CGCTCGACCAAGCGGAGGCGCTCTCTGGTCAATGTTGGAAATCTTTCGAAGTTTGGAAA	3047
QY	1098	TCTCTCCAACTACCTCAAGAGGAAAAGTGACTATATTTTCTGTCAACAGATCCAGCACT	1157
Db	3048	CTATCACTTACTTACGGGGCAGAGAATGAAATTTGTTCCCTATAGAGCAAGGGGCG	3107
QY	1158	ACACATGGAGGCTAAGAAAGAAAAATGGAGCAGCGCTTGGAAACAGGCAAGAAACCAG	1217
Db	3108	ACGCTTCCGCCCAAGGCAAGCACTACGTTTGGG--GAGCTCTCGGTGATCGAAAAGACG	3164
QY	1218	ACTAGATAGCTACACAGCAGGAGAAAGCTTTTGGAGCTCCGSGCTTTCAGAGAGATTAAG	1277
Db	3165	CTGGACAGCATCACACAGCAGCAGAGCTCTGCAGCTCAGGCTTTGTGTGAGGAAATC	3224
QY	1278	TCTGAGTGTGTTGAGGAAGAGAGGATTTCTGACGGTTTCTACAAGAGCCCATCACTAT	1337
Db	3225	GCTCAGTGTGTAGAGGAAGAAAGAGCTTCTGAAGAACTGTACAAGCACTTCAGCCTT	3294
QY	1338	GGAAGATCTGATTTCTACAGTTTTCAGTGGCCAGAGCATGGNGTTTCTGTCTTCAG	1397
Db	3285	GGAGCTATCTATCTGTACAGCTTCCAGTGGCTAAGGCGATGGAGTCTTGSCATCAAG	3344
QY	1398	AAAGTGCAATCATCGGACCTTGGCAGCGAGAAACATCTTTTATCTCAGAAACAGTGCTG	1457
Db	3345	GAGGTGTATCCACAGGACCTGGCAGCAGCAACATCTCTATCGGAGAGAACTGCTG	3404
QY	1458	GAAATTTGTGATTTGGCTTTCGCGGGATATTTAAGAACCCGATTAATGTGAGAAA	1517
Db	3405	TAAGATCTGTGACTTGGCTTGGCCGGGCAATTAAGAACCCGATTAATGTGAGAAA	3464
QY	1518	AGGAGATCTGCACCTTCTCTGAATGGATGGCTCCGCAATCTATCTTCACAAATCTA	1577
Db	3465	AGGAGATGCCAGCTCTCTTGAAGTGGATGGCCCGGAAACCAATTTTGCAGAGTATA	3524
QY	1578	CAGCACAAGACGACGTTGGTCTTACGAGTATCTGTGTGGGAAATCTTCTCTAGG	1637
Db	3525	CACAATTCAGACGAGTGTGTGTTTTCGGTGTGTTCTCTGGGAAATATTTCTCTAGG	3584
QY	1638	TGGTCTCCATACCCAGAGTACAANTGGATGAGGACITTTTCAGTCCCTGAGCGAAGG	1697
Db	3585	TGCTCCCATACCTCGGGGTCAAGATGTGAGAGAAATTTGTAGGAGATTTGAAGAAG	3644
QY	1698	CATGAGGATCAGAGCTCTCTGAGTACTCTACTCTGAAATCTATCAGATCATCTCGACTG	1757
Db	3645	AACTAGAATCGGGCTCTCTCACTACACTACCCAGAAATGTACCAGACCATCTCGACTG	3704
QY	1758	CTGSCACAGAGACCCAAAGAGAGGCCAAGATTTGAGAACTTTCGAAAACACTAGTGA	1817
Db	3705	CTGGCATGAGACCCCAACAGAGACCTCTGTTTCAGAGTTGTGTGAGCACTTTGGGAA	3764
QY	1818	TTTGCTTCAAGCAATGTACAAGCAGATGGTAAGACTACAT	1859
Db	3765	CTCTCTGCAAGCAATGCGCAGCAGGATGGCAAGACTATAI	3805

RESULT 7

US-07-813-593-3

: Sequence 3, Application US/07813593

: Patent No. 5185438

: GENERAL INFORMATION:

: APPLICANT: Lemischka, Ibor R.

: TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

: NUMBER OF SEQUENCES: 4

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: IMPLONE SYSTEMS INCORPORATED

: STREET: 180 VARICK STREET

: CITY: NEW YORK

: STATE: NEW YORK

: COUNTRY: U.S.A.

: ZIP: 10014

Db 2490 GAGAGTCAATATCTCGCGGACCTGAGGATGATGCGATGTTCTTCTGCTCCCTTCTGT 2549
QY 681 CCTCTTATCCAAATGAAAGGTC---TCTCTGAAATAAGAGACTGACTACCTATC 737
Db 2550 CATCTCGTACGAGCGGTTAAGCGGCAATGAAGGGAACAGACAGACGACCTATGTC 2609
QY 738 AATTAATAGGACGAGATGAGATGCTTTGGATGAGCAGTGTGAGCGCTTCTCTTAAGA 797
Db 2610 TATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2669
QY 798 TCCAGCAAGTGGGAGTTTCCCGGAGAGACTTAACTGGGCAAACTACCTTGAAGAG 857
Db 2670 TCCAGCAAGTGGGAAATCCCGAGGACCGGCTGAACATAGGAAACCTTGGCCCGG 2729
QY 858 GCTTTTGGAAAGTGGTTCAGCAATCAGCAATTTGGCATTAACAAATCACTTCTCCG 917
Db 2730 TCCCTTGGGCAAGTATGAGGAGAGGCTTTTGGAAATGAAAGAGAGGAGCTTGA 2789
QY 918 GACTTGGCTGTGAATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
Db 2790 AACAGTACGCTGAGATGTTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2849
QY 978 GACTGAGCTAAATCTTGAACCAATTTGGCCACATCTGAGGAGGAGGAGGAGGAGGAG 1037
Db 2850 GTCTGACTCAAGATCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2909
QY 1038 ACCCTGACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097
Db 2910 CCGCTGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2969
QY 1098 TCTCTCAACTACTTCAAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1157
Db 2970 CTTATCACTTACTTACGGGCAAGAGAAATGAAATTTCTCTATAGAGCAAGGAGG 3029
QY 1158 ACATATGAGGCTTAAGAAAGAAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1217
Db 3030 AGCTTTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3086
QY 1218 ACTAGATAGGCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1277
Db 3087 CTGGAGAGCAATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3146
QY 1278 TCTGAGTATGTTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1337
Db 3147 GCTCAGTATGATAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3206
QY 1338 GAGAGATCTGATTTTACAGTTTTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1397
Db 3207 GAGAGATCTGATTTTACAGTTTTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3266
QY 1398 AAGTGTGATTTTACAGTTTTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1457
Db 3267 GAGTGTATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3326
QY 1458 GAGATTTGATTTTGGGCTTCCCGGAGATTTTAAAGAGGAGGAGGAGGAGGAGGAGGAG 1517
Db 3327 TAAGATCTGATTTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3386
QY 1518 AGGAGATCTGATTTTGGGCTTCCCGGAGATTTTAAAGAGGAGGAGGAGGAGGAGGAG 1577
Db 3387 AGGAGATCTGATTTTGGGCTTCCCGGAGATTTTAAAGAGGAGGAGGAGGAGGAGGAG 3446
QY 1578 GAGCAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1637
Db 3447 CACATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3506
QY 1638 TGGGCTTCCATACCGAGGAGTAAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1697
Db 3507 TGGCTTCCCATACCGAGGAGTAAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3566
QY 1698 CATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1757

Db 3567 AACTAGAAATGGGCTCTGACTACACTACCCAGAAATGTACCAGACCTCTGGAGCTG 3626
QY 1758 TTGGCAGAGAGCCCAAGAAAGAGCCCAAGATTTGCAGAACTTTGGAAACTAGGTGA 1817
Db 3627 CTGSCATGAGGAGCCCAAGCAACAGAGACCTCTCTTTTCAGATTTGGTGGAGCATTTGGGAAA 3686
QY 1818 TTGCTTCAGCAAAATGTACACAGATGTACACAGATGTAAAGACTATAT 1859
Db 3687 CTTCTGCAAGCAATGCGCAGCAGATGCGCAAGACTATAT 3728
RESULT 8
US-07-977-451-5
: Sequence 5, Application US/07977451
: Patent No. 5270458
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ImClone Systems Incorporated
: STREET: 180 Varick Street
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10014
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/977,451
: FILING DATE: 19921119
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US UNASSIGNED
: FILING DATE: 12-NOV-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/906,397
: FILING DATE: 26-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCI/US92/05401
: FILING DATE: 26-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: TW 81102961
: FILING DATE: 15-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCI/US92/02750
: FILING DATE: 02-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/813,593
: FILING DATE: 24-DEC-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/793,065
: FILING DATE: 15-NOV-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/728,913
: FILING DATE: 28-JUN-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/679,666
: FILING DATE: 02-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Feit, Irving N.
: REGISTRATION NUMBER: 28,601
: REFERENCE/DOCKET NUMBER: LEM-3-7P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5406 base pairs

STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: FICOPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,507
FILING DATE: 19920817
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/579,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
NAME/KEY: mat_peptide
LOCATION: 208..4308
US-07-946-507-3

Query Match 33.7%; Score 708.4; DB 1; Length 5406;
Best Local Similarity 65.0%; Pred. No. 4.9e-210;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;
QY 201 CACTTTTACCATCATGATTTTCCTGCGAGATTTCAGGACCTATGCGTGCAG 250
DB 2070 CACTTTGATTTGCGATTTTCAGATTCCTCTCTGAGGACGAGTATGTTGCTC 2129
QY 261 AGCCAGGAGTATACACAGGAGGAGAAATCTCCAGAGAAAGAAATTAACATCAGAGA 320
DB 2130 TCTCAAGATAG 2189
QY 321 TCAG 380
DB 2190 GCGGATGCGAG 2249
QY 381 CACCAGTTCAGATTCATGATTCATGATTCATGATTCATGATTCATGATTCATGATTC 440
DB 2250 CATTGAAGTACATTCATGATTCATGATTCATGATTCATGATTCATGATTCATGATTC 2309
QY 441 CAACACAG 500
DB 2310 CAACAG 2369
QY 501 TATTGAAG 560

DB 2370 TATCCGACAGGTTGAG 2420
QY 561 GCGCTCTGTGAG 620
DB 2430 TGGCTGTGAG 2480
QY 621 GAGCTGTGAG 680
DB 2490 GGAAGTCAATTCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2549
QY 681 CCGCTTATCGGAG 737
DB 2550 CATTCGAG 2609
QY 738 AATTATATGAG 797
DB 2610 TATTGTCAATTCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2669
QY 798 TGGCAG 857
DB 2670 TGGCAG 2729
QY 858 GCGCTTGTGAG 917
DB 2730 TGGCTGTGAG 2780
QY 918 GACTGTGCTGTGAG 977
DB 2790 AACAGTACAG 2840
QY 978 GACTGTGCTGTGAG 1037
DB 2950 GCTGTGAG 2909
QY 1038 AGCTGTGCTGTGAG 1097
DB 2910 GCGCTGTGAG 2969
QY 1098 TCTGTGAG 1157
DB 2970 CCAATCACTTACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3020
QY 1158 ACAGATGAG 1217
DB 3030 AGCTTTCGCGAG 3086
QY 1218 ACTGTGAG 1277
DB 3087 CTGGAG 3146
QY 1278 TCTGTGAG 1337
DB 3147 GCTGTGAG 3205
QY 1338 GGAAGTCAATTCCTGCGAG 1397
DB 3207 GGAAGTCAATTCCTGCGAG 3266
QY 1398 AAGTGTGAG 1457
DB 3267 GAAGTGTGAG 3326
QY 1458 GAAGTGTGAG 1517
DB 3327 TATGTGTGAG 3386
QY 1518 AGAGATGAG 1577
DB 3387 AGAGATGAG 3446
QY 1578 CAGCAG 1637

Db 3447 CACAAITCAGACGAGTGTGGCTTTCGGTGTGTCTCTGGAAATATTTCTTAGG 3506
QY 1638 TGGGTCTCCATACCCAGGAGTACAAATGGATGAGGACTTTTCCAGTCCCTGAGGGAAGG 1697
Db 3507 TGGTCTCCATACCCCTGGGTCAAGATTCATGAAGAAATTTTGTAGGAGATTTGAAGAGG 3566
QY 1698 CATGAGATGAGGCTGTGAGTACTTCTCTGAAATCTATCAGATCATCTGGAGTGT 1757
Db 3567 AACTAGAAATGGGGTCTCTGACTACTACCTCCAGAAATGTACGACCACTGTGGAGTGT 3626
QY 1758 CTGGCAGACAGACCCCAAGAGGAGCCAGATTTTCAGAACTTGTGAAAACCTAGGTGA 1817
Db 3627 CTGGCAGACAGACCCCAAGAGGAGCCCTGTTTTCAGATTTGTGGAGCAATTTGGGAA 3686
QY 1818 TTGCTCTCAAGCAATGTACAAAGGATGTGTAAGACTACAT 1859
Db 3687 CCTCTGCAAGCAATGCGCCAGGAGGATGCAAGACTATAT 3728

RESULT 10

-08-252-517-5
Sequence 5, Application US/09252517
Patent No. 5548065
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: IOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,517
FILING DATE: 31-OCT-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 25-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCI/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCI/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.

REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
US-08-252-517-5

Query Match 33.7%; Score 708.4; DB 1; Length 5406;
Best Local Similarity 65.0%; Pred. No. 4.9e-210;
Matches 1080; Conservative 0; Mismatches 576; Indels 2;

QY 201 CACTTAACTTACCATCATGAATGTTTCCCTGCAAGATTCNGSCACCTATCGCTGCAG 260
DB 2070 CATCTTGAITGIGCAATTCAGAAIGCTCTCTGCAAGCAAGGCGACTATGTTTCTC 2129
QY 261 AGCCAGCAATGTATACACAGGGGAGAAATCTCTCCAGAGAAAGAAATTTACAATCAGAG 320
DB 2130 TGTCAAGATAGAGACCAAGAAAGACATTCCTTGGTCAACAGCTCATCTCTCAGA 2189
QY 321 TCAGGAGACCATACCTCTTCCGAAACCTCAGTNTACACAGTGGCCATCAGAGTTC 380
DB 2190 GCGATGGCACCCTCATGATCCCGGAAATCTGGAGAAATCAGACAAACCAATGGCGAGAC 2249
QY 381 CAGCAGTTTAGACTGTCTAATGTTGTTCCCGAGGCTCAGATCATCTTGGTTTAAAAA 440
DB 2250 CATTTGAAGTGAATTCCTCCAGCACTGGAATCTTACCCCAACATACATGTTTCAAGA 2309
QY 441 CAACCAAAATACAAAGAGAGCTGGAATTTTAGACCAAGAGAGAGAGAGCTGT 500
DB 2310 CAACGAGACCTCTGGTAGAAGATTCAGGCAATTTGCTAGAGATGGGAACCGGAACCTGAC 2369
QY 501 TATTGAAGAGTACAGAGAGAGATGAAGGTGTCTATCACTGCAAGCCCAACCAAGAA 560
DB 2370 TATCCGAGGTTGAGGAAGAGAGATGGAGGCTCTACACCTGCGAGGCTTGAATGTCT 2429
QY 561 GGGCTCTGTGGAAGTTTCAAGTATCTTCAAGCAACCTCTGGAAGCAAGTCTAATCT 620
DB 2430 TGGCTGTGAAGAGGAGAGCTCTTCAATAAGAGGTGCCAGGAAAGAACCAACTT 2489
QY 521 GGAGCTCATCTCTAATCATGCACTGTGTGGTGGAGCTCTCTTCTGGCTCTATTAC 680
DB 2490 GGAAGTATTTCTCTGCGGCACTGAGTGTGCAATGTTCTTCTGCTCTCTCT 2549
QY 681 CCTCTTATCCGAAATGAAAAGTCTCTCTCTGAAATTAAGAGTCTGACTACCTATC 737
DB 2550 CATTCGTACGAGCCGTTAAGCGGCAATGAAGGGAACCTGAAGACAGGCTACTTGT 2609
QY 738 AATTAATGGACCCAGATGAATTCCTTTGGATGAGAGTGTGAGCGGCTCCCTTATGA 797
DB 2610 TATTGTATGATCCAGATGAATTCCTTGGATGAGGCTCTGAGGCTTGTCTTATGA 2669
QY 798 TGCAGCAAGTGGAGTTTGGCGGAGAGACTTAACTGGGCAAACTCACTTGGAGAGG 857
DB 2670 TGCAGCAAGTGGAGTTTCCCGAGGAGGCTGAGTGAAGTCTTGGCGGCGG 2729

Db 2190 GCGATGGCACCACATGATCAGCGGAAATCGGAGAAATCAGACAAACCAATGGCGAGAC 2249
Qy 381 CACACCTTTAGACTGTCAGTAACTGAGTGGTCCCGAGGCTCAGATCAGTCTGGTTTAAAAA 440
Db 2250 CAITGAAGTACCTTCTCAGCACTGGAATCTCAACCAACATTCATCGTTCAGAA 2309
Qy 441 CACCAACAAATACACAGAGGCTGGAATTTATAGACAGGAGAGACAGCTGTT 500
Db 2310 CAGGAGACCTTGGTAGAGATTCAGGCAITGTACTGAGATGGGACCGGAACTGAC 2369
Qy 501 TATTAAGAGATCAGAGAGAGGAGTGTCTATCAGTGCAGAGCCACCAACCAAGAA 560
Db 2370 TATCCGAGGGTGGAGAGAGGATGAGGCTCTACACCTGCCAGGCTCGAATGCT 2429
Qy 561 GGGCTGTGGAAAGTTCAGCATACCTCAGTGTCAAGACCTCGGACAGTCTAATCT 620
Db 2430 TGGCTGTGCAAGAGCGGAGAGGCTCTTCAATAGAGAGTGGCCAGGAAAGACCACT 2489
Qy 621 GGAGCTGATCAGTCTAACATCAGCTGTGGCTGGAGCTCTCTCTGGCTCTCTATAC 580
Db 2490 GGAAGTATATCTCTGCTGGCAGTGTGAGTGTGCAATGCTCTCTGGCTCTCTGT 2549
Qy 691 CTTCTTCCGAAATATGAAAGGTC---TTCCTTGAATTAAGAGTACTGACTATC 737
Db 2550 CAITCTCGTACGACCGTTAAGCGGCGCAATGAAGGGAATGAAGACAGCTTCTGTC 2609
Qy 738 AATTAATAGGACCCAGATGAAGTTCCTTTGAGTGAAGTGTGAGGCTCTCTATGA 797
Db 2610 TATGATGATGATCCAGATGAATTCGCTTGAAGTGTGAGGCTGTGAGGCTTGTCTATGA 2659
Qy 798 TCCAGCAGTGGGAGTTCGCCGGAGAGACTTAACTGGGCAATCAGTGGAGAGG 857
Db 2670 TGGCAGAGTGGAAATCCCGAGGACCGCTGAACTAGAAACCTCTTGGCGGCG 2729
Qy 858 GGGTTTGAAGAGTGTTCAGCATCAGCAATTTGGCAATTAAGAACTACCTACCTGCGG 917
Db 2730 TGGCTTGGCCAGTGTGAGGAGGAGGCTTTTGAATTCAGACAGGAGCTTGCAT 2789
Qy 918 GACTGTGCTGTGAAATGCTGAAGAGGGGCGACGGCCAGGAGTACAAAGCTGTAT 977
Db 2790 AACAGTAGGCTCAGATGTGAAGAGGAGGACACAGGAGATCGAGCCCTCAT 2849
Qy 978 GACTGAGCTAAATCTTACCCACATTTGCCACCATCTGACGTGGTTAACCTGTGG 1037
Db 2850 GTCTGAATCAAGATCTCTATCCATGTTGGTACCACTCAATGTGGAGACCTCTGCG 2909
Qy 1038 AGCTCCACAGCAAGAGGCTCTGATGGTGTATGTTGAATCTGAAATAGCAAA 1097
Db 2910 CGCTGCACCAAGCGGAGGCGCTCTCATGTTGATTTGGAATCTCGAAGTTTGA 2969
Qy 1098 TCTCTCAACTACCTCAAGAGCAACGTGACTTATTTTCTCAACAGGATGCACTACT 1157
Db 2970 CCTATCAACTTACTACGGGCGAGAGAAATGAATTTGCTTATAGAGCAAGGCGC 3029
Qy 1158 ACATGAGGCTTAAAGAAAGAAATGAGCGGCTCTGCAACAGGCAAGAAACCAAG 1217
Db 3030 AGCTTCCGCGAGGCAAGGACTACGTGGG---GAGCTCTCCGTGGATCTGAAAGAGC 3086
Qy 1218 ACTAGATAGCTCAGCAGCAGCAAGCTTTGGAGCTCCGCTTTCAGGAGATTAAG 1277
Db 3087 CTGGCAGGATCAGCAGCAGCAGCTCTGCCAGCTCAGCTTTGTTGAGGAGAACT 3146
Qy 1278 TCTGAGTGTGTGAGGAGAGGAGGATTTCTGACGGTTTCTCAAGGAGGCGCACTAT 1337
Db 3147 GCTCAGTGTGTGAGGAGAGGAGGCTTCTGAAGAGCTGTACAGGACTTCTGACCT 3206
Qy 1338 GGAAGATCTGATTTCTTACAGTTTTCAGTGGCCAGGAGGATGAGTTCTGCTTCCAG 1397
Db 3207 GGAGCATCTCATCTGTACAGCTTCCAAAGTGGCTAAGGCGATGGAGTTCTTGGCATCAAG 3266
Qy 1398 AAGTGTATCTATGAGGAGCTTGGAGGAGGAGAACTTCTTCTGAGGAGCAAGCTGT 1457
Db 3267 GAAGTGTATCCAGAGGAGCTTGGCAGCAGCAACAACTTCTCTATCGGAGAGAGTGTGT 3326

RESULT 12

US-08-601-891-5

; Sequence S, Application US/08601891

; Patent No. 5747651

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Imclone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/08/601,891

; FILING DATE: 15-FEB-1996

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/977,451

; FILING DATE: 19-NOV-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/906,397

; FILING DATE: 26-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US PCT/US92/05401

; FILING DATE: 26-JUN-1992

; APPLICATION NUMBER: TW 81102961

; FILING DATE: 15-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US PCT/US92/02750

; FILING DATE: 02-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/813,593

; FILING DATE: 24-DEC-1991

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/793,065
 FILING DATE: 15-NOV-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/728,913
 FILING DATE: 28-JUN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/579,666
 FILING DATE: 02-APR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Felt, Irving N.
 REGISTRATION NUMBER: 28,601
 REFERENCE/DOCKET NUMBER: LEM-3-7P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-645-1405
 TELEFAX: 212-645-2054
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5406 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 208..4311
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 265..4308
 FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: 208..264
 US-08-601-891-5

Query Match 33.7%; Score 708.4; DB 1; Length 5406;

Best Local Similarity 55.0%; Pred. No. 4.9e-210; Matches 1080; Conservative 0; Mismatches 576; Indels 2; Gaps 2;

Oy 201 CACTCTTAATCTTACCATCAATGATTTCCCTGCAAGATTCAGGACCACTATGCTGCGAG 260
 Db 2070 CATCTTGATTTGGCAATTCAGATGCCCTCTGCGAGGACCAAGCGGACATGTTGCTC 2129
 Oy 261 AGCCAGGAATGTATACAGAGGAGAGAAATCTCCAGAGNAGNAGNAGNAGNAGNAGNAG 320
 Db 2130 TGCTCAAGATAAGAGAGACCAAGAAAGACATTCCTGCTCAAGACGCTCATCACTAGA 2189
 Oy 321 TCAGGAAGCACCATACCTCTCGGAAACCTCAGTCATCACACAGTCGCGCATCAGGAGTTC 380
 Db 2190 GCGCATGGCACCATGATCACCAGGAATCTGGAGNATCAGACACACCAATTCGAGNAC 2249
 Oy 381 CACCCTTTAGACTGTATGCTAATGGTGTCCCGAGGCTCAGATACATCTGGTTTAAAAA 440
 Db 2250 CATTTGAAGTGACTTGGCCACGATCTGGAATCTACCCGACACACATACATGTTTCAAAGA 2309
 Oy 441 CRACCAAAATACACAAAGAGCTGGAATTAATTCAGCCAGAGAGAGAGAGAGAGAGTGT 500
 Db 2310 CAACGAGACCTTGGTAGAAGATTCAGGCATGTACTAGAGATGGGAACCGGACCTGAC 2349
 Oy 501 TATTGAAGAGCTACAGAAAGAGTGAAGTGTCTATCACTGCAAGAGCCACCAACCAAGAA 560
 Db 2370 TATCCGAGGGTGCAGAGAGAGATGGAGGCTCTACACCTGCGAGGCTGCAATGCTT 2429
 Oy 561 GGGCTCTGTGAAAGTTTCAGATACCTCAGTGTTCAGAGAACCTCGGACAGATCTATCT 620
 Db 2430 TGGCTGTGAAGAGCGGAGAGCTCTTCATATAGAAGGTGCCGAGAAAGACCAAT 2489
 Oy 621 GGAGCTGATCATCTAACATCAGCTGTGGCTGGAGCTCTCTCTCTGGCTCCCTTAC 680
 Db 2490 GGAAGTCATATCTCTCGTGGCAGTCAGTGATTCGCAATGTTCTTGGCTCTCTTCT 2549

Oy 681 CCTCTTAATCCGAAAAATGAAAGGTC---TCTTCTCAAAATAAGAGCTGACTACCTATC 737
 Db 2550 CATCTCTGACGACCGTTAAGCGGCCAATGAAGGGGAAGTGAAGACAGAGGCTACTTGTC 2609
 Oy 738 AATTATATGACCCAGATGATGTTCTTTGATGACAGCTGTGACGCGCTCCCTTATGA 797
 Db 2610 TATGTGATGATCCAGATGAATGCCCTTGGATGAGGCTGTGAAGCTTCCCTTATGA 2669
 Oy 798 TCCAGCAAGTGGAGTTTCCCGGAGAGACTTAAACTGGCAAAATCACTTTGGAAGAG 857
 Db 2670 TCCAGCAAGTGGAGTTTCCCGGAGAGACTTAAACTGGCAAAATCACTTTGGAAGAG 2729
 Oy 858 GCGTTTGGAAAAGTGTTCAGCATCAGCATTTGGCATTAAGAAATCAGCTACGTCGG 917
 Db 2730 TCCCTTGGCAAGTGTTCAGCATCAGCATTTGGCATTAAGAAATCAGCTACGTCGG 2789
 Oy 918 GACTGTGGCTGTGAAAATGCTGAAAGAGGGGCCACGCCAGGAGTACAAAGCTCTGAT 977
 Db 2790 AACAGTACCGGTCAAGATGTTGAAAGAGGAGCAACACAGCGGAGCATCGAGCCCTCAT 2849
 Oy 978 GACTGAGCTAAAATCTTGACCCACATTTGGCCACATCTGAAGTGTGTTAAGCTGCTGG 1037
 Db 2850 GTCTGAATCAAGATCTCTATCCACATTTGGTCAACATCTAAATGTGTGAACCTCTTAG 2909
 Oy 1038 AGCTGCAAGCAAGAGGAGGCTCTGATGGTGTGATTTGAATGCTGCAAAATATGAAA 1057
 Db 2910 AGCTGCAAGCAAGAGGAGGCTCTGATGGTGTGATTTGAATGCTGCAAAATATGAAA 2969
 Oy 1098 TCTCTCAACTACCTCAAGAGCAAGGCTGACTTATTTTCTCAAGAGGATCGACACT 1157
 Db 2970 CCAATCACTTACTTACGGGCAAGAGAAATCAATTTGTTCCCTATAAGAGCAAGAGG 3029
 Oy 1158 ACATATGAGCTTAAGAAAGAAATGAGCGGCTGGAACAGGCAAGAAACCAAG 1217
 Db 3030 AGCTTCCGAGGCAAGGAGTACGTTGGG---GAGCTCTCGGTGATCTGAAAGAGC 3086
 Oy 1218 ACTAGATAGCTCACAGCAGCGAAAGCTTTCCGAGCTCCGGCTTTCCAGGAAGATAAAG 1277
 Db 3087 CTGAGCAGCATACAGAGCAGGAGCTCTGCGAGCTCAGGCTTTGTTGAGGAGAAATC 3146
 Oy 1278 TCTGAGTGTATTTGAGGAGGAGGATTTCTGCGGTTTCTTACAGGAGCCCATCACTAT 1337
 Db 3147 GCTCAGTGTATGAGGAGGAGGATTTCTCAAGAACTGTACAAAGACTTCTCTGACCTT 3206
 Oy 1338 GGAGATCTGATTTTACAGTTTCAAGTGGCCAGAGGCTGAGTTCTCTGCTCTCCAG 1397
 Db 3207 GGAGATCTGATTTTACAGTTTCAAGTGGCCAGAGGCTGAGTTCTCTGCTCTCCAG 3266
 Oy 1398 AAGATCTGATTTGAGGAGGAGGATTTCTGCGGTTTCTTACAGGAGCCCATCACTAT 1457
 Db 3267 GAAGTGTATCCAGGAGGACTTGCAGCAGCAAGAACTTCTCTCTGAGGAGAAATGTGT 3326
 Oy 1458 GAAGATTTGATTTTGGCTTGGCGGAGTATTTATAGAACCCGATTTATGTGAGAA 1517
 Db 3327 TAAATCTGTGATTTGGCTTGGCGGAGTATTTATAGAACCCGATTTATGTGAGAA 3386
 Oy 1518 AGGAGTACTCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1577
 Db 3387 AGGAGTACTCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3446
 Oy 1578 CAGCAGCAAGAGGAGTGTGTTCTTACGAGTATTTGCTGTGGGAAATCTTCTCTTAG 1637
 Db 3447 CACAATTCAGAGGAGTGTGTTCTTACGAGTATTTGCTGTGGGAAATCTTCTCTTAG 3506
 Oy 1638 TGGGCTCTCATACCCAGGAGTACAAATGATGAGGAGTATTTGAGTGGCTGAGGAGAG 1697
 Db 3507 TGGCTCTCATACCCAGGAGTACAAATGATGAGGAGTATTTGAGTGGCTGAGGAGAG 3566
 Oy 1698 CATGAGGATGAGGAGTCTCTGAGTACTCTACTCTGAAATCTATCAGATCACTGAGTGT 1757
 Db 3567 AACTAGAAATGCGGGCTCTCTGAGTACTCTACTCTGAAATCTATCAGATCACTGAGTGT 3626
 Oy 1758 CTGSCACAGAGACCCCAAGAAAGAGGCAAGATTTGAGAACTTTGTGAAAACTAGTGA 1817

Db 3627 CTGGCAAGGACCCACACAGACCTCGTTTCAGAGTGTGGAGCAATTTGGAAA 3686
Qy 1818 TTTCCTTCAAGCAATGTACACAGGATGTAAGACTACAT 1859
Db 3687 CCTCTGCAAGCAATGTACACAGGATGTAAGACTACAT 3728

RESULT 13

US-09-021-324-5

Sequence 5, Application US/09021324

Patent No. 5912133

GENERAL INFORMATION:

APPLICANT: Lemischka, Ibor R.

TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: ImClone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/021,324

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/977,451

FILING DATE: 1992-11-19

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Felt, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 5406 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

Qy 201 CAGCTTATCTTACCATCATGATGTTTCCCTGCAAGATTTCAGGACCTATCCCTGAG 260
Db 2070 CATCTTATGTTGGCAATTCAGAAATGCCCTCTCTGCGAGCAAGGCGACTATGTTGCTC 2129
Qy 261 AGCCAGGAAATGTATACACAGGGGAGAAATCCCTCCAGAGAGAAATTTACAATCAGAGA 320
Db 2130 TGCCTAAGATAGAGACCAAGAAAGACATTCCTGGTCAACAGCTCATCATCTAGA 2189
Qy 321 TCAGGAGACCATACCTTCCTCGGAACCTCAGTGATCACACAGTGGCCATCAGAGTTC 380
Db 2190 GCGATCGGCCCATGATCATCGCGAAATCTGGAGAAATCGGAGAAACACCAACCTTGGCGAGAC 2249
Qy 381 CACCACCTTATGACTGTCATGCTAATGGTGTCCCGAGGCTCAGATCAGTGTGTTTAAAGA 440
Db 2250 CATGAGTGCATTCCTCCAGCATCTGGAATCTACCCACACATTTACATGGTTCAAGA 2309
Qy 441 CAACCAAAATACAAAGAGCCCTGGAAATTTTATAGCAGGAGAGACAGCAGCTGTT 500
Db 2310 CAAGAGACCTGGTAGAAGATTTCAGGCATTGTACAGAGATGGGAACCGGAACCTGAC 2369
Qy 501 TATGAAGATCACAGAGAGGATGAAGTGTCTATCATCTGCAAGCCACCAACAGAA 560
Db 2370 TATCCGAGGTGAGGAGAGGAGTGGAGGCTCTACACCTGCCAGGCTGCAATGTCT 2429
Qy 561 GGCCTGTGGAAGTTTCAGCATACCTCAGTCTCAAGAACTCCGACCAAGTCTAATCT 620
Db 2430 TGGCTGCAAGAGCGGAGACGCTCTTCATATAGAGGTGCCAGGAAAGACCACTT 2489
Qy 621 GGAGCTGATCCTTAACATGCACTGTGGCTGGAGTCTCTTCTTGGCTCCTATTAC 680
Db 2490 GGAAGTATATCTCTCGTGGCAGTGTGATGCTGATGCTGCTGCTGCTGCTTCTTGT 2549
Qy 681 CCTCCTATCCGAAATGAAAGSTC---TCTTCTGAATAGAGTCACTACCTATC 737
Db 2550 CATTCGTACGACCCGTTAAGCGGGCCAAAGAGGGGAACTGAGAGACAGGCTACTTGT 2609
Qy 738 AATATATGAGCCAGGATGAAGTTCCTTTGATGAGCAGTGTGAGCGGCTCCCTTATGA 797
Db 2610 TATGTCTGATCCAGATGAATTTGCCCTTTGGATGAGCGCTGTGAACGCTTGCCTTATCA 2669
Qy 798 TGCACAGTGGAGTTCGCCGGGAGAGACTTAACTGGGCAATCAGTTCGGAAGAG 857
Db 2670 TGCAGCAAGTGGGAATTCGCCGAGGACCGGCTGAACCTTAGAGAGAGGCTTGTGCGCGG 2729
Qy 858 GGCCTTTGGAAAGTGGTTCAAGCATCAGCATTTGGCAATTTGAATAACCTACCTACGTCGG 917
Db 2730 TGCCTGCGCCAGTGTAGGCGAGAGCTTTTGGAAATGACAGACAGGCTTGCAT 2789
Qy 918 GACTGTGGCTGTGAAATGCTGAAGAGGGGGCCAGCGCCAGGAGTACAAAGCTGTAT 977
Db 2790 AACAGTAGCCGCTCAAGATGTTCAAGAGAGGAGCAACACACAGCGAGCATCGAGCCCTCAT 2849
Qy 978 GACTGAGCTTAAATCTTGACCCACATTCGCCACCACTTGACGCTGGTTACCTGTGG 1037

Db 2850 GTCTGAACCTAAGATCTTCAATCCACATTTGGTCACCAATCTCAATGTGGTGAACCTCTTAGG 2909
QY 1038 AGCTTCACCAAGCAAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1097
Db 2910 CGCTTCACCAAGCAAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 2969
QY 1098 TCCTTCACCTACTCTCAAGACCAAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1157
Db 2970 CCTATCAACTTACTTACGGGGCAAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 3029
QY 1158 ACATGTGAGGCTTAAAGAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1217
Db 3030 AGCTTCACCAAGCAAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 3086
QY 1218 ACTAGATAGGCTTACCAAGCAAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1277
Db 3087 CTTCGACGATCAACCAAGCAAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 3146
QY 1278 TCTGATGATGTTGAGGAGGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1337
Db 3147 GCTCAGTATGTTGAGGAGGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 3206
QY 1338 GGAAGATCTCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 1397
Db 3207 GGACATCTCATCTGTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 3266
QY 1398 AAGTCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 1457
Db 3267 GAAGTGTATCTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 3326
QY 1458 GAAGTGTATCTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 1517
Db 3327 TAAATCTGTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 3386
QY 1518 AGGAGATCTTCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 1577
Db 3387 AGGAGATCTTCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 3446
QY 1578 CAGCACAAGAGGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1637
Db 3447 CACAATCTACAGGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 3506
QY 1638 TGGTCTCTCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 1697
Db 3507 TGGTCTCTCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 3566
QY 1698 CATGAGATCAGAGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1757
b 3567 AACTAGATCTTCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 3626
QY 1758 CTGGCAGAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 1817
Db 3627 CTGGCAGAGAGGCTCTGTATGTGATGTTGTAATCTGCAATATGCAAA 3686
QY 1818 TTGCTTCAAGCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 1859
Db 3687 CCTCTGCAAGCAATCTTACAGTTTCAAGTGGCCAGGAGGCTCTGTATGTGATGTTGTAATCTGCAAA 3728

RESULT 14

PCT-US92-02750-7

Sequence 7, Application PC/US9202750

GENERAL INFORMATION:

APPLICANT: LEWISCHKA, ITHOR R.

TITLE OF INVENTION: Totipotent Hematopoietic Stem Cell

TITLE OF INVENTION: Receptors And their Ligands

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESS: INCUBATION SYSTEMS INCORPORATED

STREET: 180 VARICK STREET

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: US

ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02750
FILING DATE: 19920402
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FELI, IRVING N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: DEM-3-PPPT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO. 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 208..4308
PCT-US92-02750-7

Query Match 33.78; Score 708.4; DB 5; Length 5406;

Best Local Similarity 65.08; Pred. No. 4.9e-210;

Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;

QY 201 CACTCTTAATCTTACCAATCATGATGTTTCCCTGCAAGATTCAGGCACCTATGCTGAG 260
Db 2070 CATCTTGATGTGGCAATTCAGATGCTCTCTGAGGACCAAGCGGACTATGTTGCTC 2129
QY 261 AGCCAGGAATGTATACAGAGGGAAGAAATCTCTCAGAAAGAAATATACATCAGAGA 320
Db 2130 TCGTCAAGATTAAGAACCAAGAAAGACATTCCTGGTCAACAGAGCTCATCATCTGA 2189
QY 321 TCAGGAAGACCACTACTCTCTGCAAGACCTCAGTATCATCAGAGTGGCCATCAGCAG 380
Db 2190 GCGATGCGACCATCATCATCAGCGAAATCTGAGATCAGACAAACCATTCGCGAGAC 2249
QY 381 CACCACCTTTAGACTGTCTATGTTGGTGTCCCGAGCCTCAGATCAGCTTGTGTTTAA 440
Db 2250 CATTCAGATGCTTGGCCAGCATCTGGAATCTCTACCCACACATTACATGTTTCAAGA 2309
QY 441 CAACCAAAATACACAAGACCTTGGAAATTTATTAGGACCAAGACAGCAGCTTGT 500
Db 2310 CAACGAGACCTTGGTGAAGATTCAGGCAATGTACTGAGAGATGGACCGGACCTTAC 2369
QY 501 TATTCAGAGAGTACAGAAGAGGATCAAGGTCTCTATCATCTGCAAGACCAACCAACA 560
Db 2370 TATCCGAGGCTGAGGAGGAGGATGAGGCTCTACACCTGCGAGGCTTCAATGTTCT 2429
QY 561 GGGCTCTGTGGAAGTTTACAGATACCTCAGTGTTCAGGAACTTCGACAGCTATCT 620
Db 2430 TGGCTGTGCAAGAGCGGAGACCTCTTCAATATAGAGGTGCGCCAGGAAAGACCACT 2489
QY 521 GGAGCTGATCAGCTTAACTATGACCTGTGCTGGCTGCGACCTCTCTGCTGCTCTTAT 580
Db 2490 GGAAGTCAATTCCTGCTGCGACCTGCAAGTATGCAATGTTCTTCTGCTCTTCT 2549
QY 581 CCTCTTATCCGAAAAAAGAAAGTCTCTCTCTGCTGCTCTCTCTCTCTCTCTCTCT 737
Db 2550 CATCTCGTACGACCTTAAAGCGGCAATGAAGGGAAGTGAAGACAGGCTATCTCT 2609
QY 738 AATTATTAAGGACCAAGATGAAGTCTCTTGGATGAGCAGTGTGAGCGCTCCCTTGA 797

```

Db 3587 C T C T C C A A G C A A A T G C G C A G A G A T G G C A A A G A C T A T A T 3728

RESULI 15
PCI-US92-05401-5
: Sequence 5, Application PC/TUS9205401
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: TOTOPOENT HEMATOPOIETIC STEM CELL
: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
: STREET: 180 VARICK STREET
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: U.S.A.
: ZIP: 10014
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCI/US92/05401
: FILING DATE: 19920626
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Feit, Irving N.
: REGISTRATION NUMBER: 28.601
: REFERENCE/DOCKET NUMBER: LEM-3-PPPTT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5406 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 208..4311
: NAME/KEY: mat_peptide
: LOCATION: 208..4308
PCI-US92-05401-5

Query Match 33.7%; Score 708.4; DB 5; Length 5406;
Best Local Similarity 65.0%; Pred. No. 4.9e-210;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps

QY 201 CAC T C T T A A C T T A C C A T C A T G A A G T T T C C C T C A G A T T C A G C C A C C T A T G C C T G C A G 260
Db 2070 C A T C T G A T T G G C A T T C A G A T C C C T C T C T C G A G C C A A G C A C A T T G T T G C T 2129
QY 261 A G C C A G G A A T G T A T A C A C A G G G A A A A T C C T C C A A G A A A G A A A T T A C A A T C A G A G A 320
Db 2130 T G C T C A A G A T A A G A A G A C C A A G A A A A G A C A T T G C T G G T C A A A C A G C T A T C A T C T A G A 2189
QY 321 T A C G A A G A C C A C A T A C C T C T G G C G A A C C T A G T T A C A C A G A T G G C C A T C A G A C T T C 380
Db 2190 G C G A T G G C A C C A T G A T C A C C G G A A A T C T G A G A A T T C A G A C A A C A A C A A T T G G C G A G A C 2249
QY 381 C A C C A C T T T A G A C T G T A G C T A A G T G T C T C C C G A C C C T C A G A T C A C T T G G T T T A A A A A 440
Db 2250 C A T T C A G T A G C A C T T C C C A G A C A T C T G A A A T C C T A C C C A C A C A T T A C A T G T T C A A G A 2309
QY 441 C A A C C A A A A T T A C A A A G A G C C T G A A T T A T T T A G A C A C G A A G A C A G C A C G C T G T T 500
Db 2310 C A A C C A C C C C T G T A A G A A T T C A G C A C A T T G T A C T G A G A G A T T G G A A C C G A C C T G A C 2369

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:50:37 ; Search time 422.586 Seconds
(without alignments)
11193.756 Million cell updates/sec

Title: x51602_COPY_1900_4000
Perfect score: 2101
Sequence: 1 TATATCAGAGTGGCAA.....CGACAGCAGCAGCTGTTGG 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
tal number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAI:*

2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAI:*

3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAI:*

4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAI:*

5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAI:*

6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAI:*

7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAI:*

8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAI:*

9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAI:*

10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAI:*

11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAI:*

12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAI:*

13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAI:*

14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAI:*

15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAI:*

16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAI:*

17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAI:*

18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAI:*

19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAI:*

20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAI:*

21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAI:*

22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAI:*

23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAI:*

24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAI:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	100.0	7680	23	AA570286
2	2097.8	99.8	4017	24	ABL91752
3	715.4	34.1	4071	20	AAV99829
4	713.8	34.0	4044	24	ABL91751
5	713.8	34.0	4071	24	ABL91754
6	713.8	34.0	4225	22	AAF83308
7	710.6	33.8	4071	13	AAQ28272
8	710.6	33.8	4236	19	AAV34753
9	708.4	33.7	5404	17	AA138735

10	708.4	33.7	5404	20	AA577516	Murine flk-1 cDNA.
11	708.4	33.7	5406	14	AAQ53504	Murine flk-1 cDNA.
12	708.4	33.7	5406	14	AAQ35251	Murine flk-1 coding
13	708.4	33.7	5406	14	AAQ40916	Murine flk-1 cDNA.
14	708.4	33.7	5406	16	AAQ81014	Flk1 receptor prot
15	708.4	33.7	5406	16	AAQ79070	Mouse flk-1 cDNA.
16	708.4	33.7	5406	18	AA172119	Murine flk-1 recep
17	708.4	33.7	5470	13	AAQ64049	Sequence of murine
18	706.8	33.6	5406	13	AAQ29957	flk-1 cDNA sequenc
19	690	32.8	2352	15	AAQ74275	Truncated FLT (sVE
20	690	32.8	2352	19	AAV09333	Soluble truncated
21	690	32.8	2352	20	AAQ04435	Human soluble vasc
22	624	29.7	2523	18	AAV01457	Human VEGF recepto
23	624	29.7	2523	18	AA52101	cDNA encoding amin
24	609.6	29.0	3009	22	AA911131	Flt-1extraFasM/cy
25	541.6	25.8	3897	24	AB191753	Human polynucleoti
26	541.6	25.8	4111	21	AAQ52405	Human Flt4/VEGFR-3
27	541.6	25.8	4111	22	AAQ68952	Human Flt4/VEGFR-3
28	541.6	25.8	4195	17	AA112068	FLT4 receptor tyro
29	541.6	25.8	4195	21	AA37815	Human Flt4 recepto
30	541.6	25.8	4195	21	AA52333	Human tyrosine kin
31	541.6	25.8	4425	16	AA103090	Protein tyrosine-k
32	541.6	25.8	4795	21	AA37816	Human Flt4 recepto
33	541.6	25.8	4795	21	AA52334	Human tyrosine kin
34	541.6	25.8	9108	16	AA103104	Plasmid pRK5 tk1-1
35	540	25.7	4450	21	AAQ52210	Nucleotide sequenc
36	321.2	15.3	2313	19	AAV19347	Human soluble VEGF
37	321.2	15.3	2313	19	AAV09330	Soluble VEGF recep
38	321.2	15.3	2651	15	AAQ74268	SVGF-Ri gene. Ho
39	321.2	15.3	2651	20	AAQ4428	Human soluble vasc
40	321.2	15.3	7096	21	AA575629	Nucleotide sequenc
41	321.2	15.3	7096	24	AAQ39241	PD10-SFLT-1 vector
42	282.8	13.5	5827	14	AAQ49753	PIK gene SAL-S1.
43	282.8	13.5	5827	15	AA103101	Protein tyrosine-k
44	267.8	12.7	3069	21	AA244718	Bovine c-kit bk-1
45	259	12.3	5084	19	AAV20443	Human c-kit oncoge

ALIGNMENTS

RESULT 1
AA570286
ID AA570286 standard; cDNA: 7680 BP.
AC AA570286;
XX
XX
XX
DI 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #6090.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder: ss.
OS Homo sapiens.
XX
XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.
XX
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
FA Drmanac RT, Liu C, Tang YI;
PI
XX
XX WPI: 2001-632362/73.
DR P-PDS; ABQ5099.
XX
PI New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 1: SEQ ID No 5090; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS54197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX 50 Sequence 7680 BP: 2279 A; 1661 C; 1739 G; 2001 T; 0 other;

Query Match 100.0%; Score 2101; DB 23; Length 7680;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATATACAGATGGCCAAATGGGTTTCAATGTTACTTGGAAAAATGCCGCAAGGA 60
 DB 1900 TATATACAGATGGCCAAATGGGTTTCAATGTTACTTGGAAAAATGCCGCAAGGA 1959
 QY 61 GAGGACCTGAACCTGTCTGCACAGTTACAGTTCCTATACAGACGCTTACTTGGAT 120
 DB 1960 GAGGACCTGAACCTGTCTGCACAGTTACAGTTCCTATACAGACGCTTACTTGGAT 2019
 QY 121 TTACTGGGACAGTTAAATACAGAAATGCCATACAGTATTACGAAACAAAATGCC 180
 DB 2020 TTACTGGGACAGTTAAATACAGAAATGCCATACAGTATTACGAAACAAAATGCC 2079
 QY 181 ATCATTAGGACGACCTCCATCAGTCTTATCTTACGATGATGATGTTCCCTGCAAGAT 240
 DB 2080 ATCATTAGGACGACCTCCATCAGTCTTATCTTACGATGATGATGTTCCCTGCAAGAT 2139
 QY 241 TCAGGACCTATGCTCTGCAGAGCCAGGAATGTATACAGAGGGAAGAAATCTCCAGAG 300
 DB 2140 TCAGGACCTATGCTCTGCAGAGCCAGGAATGTATACAGAGGGAAGAAATCTCCAGAG 2199
 QY 301 AAAGAAATTAATACAGATCAGAGATCAGGACCCATACCTCTGGAAACCTCAGTGATCAC 360
 DB 2200 AAAGAAATTAATACAGATCAGAGATCAGGACCCATACCTCTGGAAACCTCAGTGATCAC 2259
 QY 361 ACAGTGGCCATCAGCAGTTCACCACTTACAGTCTCATCTTAATGCTGCTCCGAGCC 420
 DB 2260 ACAGTGGCCATCAGCAGTTCACCACTTACAGTCTCATCTTAATGCTGCTCCGAGCC 2319
 QY 421 CAGATCACTTGGTTTAAACACACCAACAAATACAGAGGCTCTGGAATTAATTAGGA 480
 DB 2320 CAGATCACTTGGTTTAAACACACCAACAAATACAGAGGCTCTGGAATTAATTAGGA 2379
 QY 481 CCAGGACGACGCTGTTTATTAAGAGATCAGAGAGATGAAGGTGCTATCAC 540
 DB 2380 CCAGGACGACGCTGTTTATTAAGAGATCAGAGAGATGAAGGTGCTATCAC 2439
 QY 541 TCGAAGGCCACCAACCAAGAGGCTCTGTGGAAGTTCAGCAATCCTACTGTTCAAGGA 600
 DB 2440 TCGAAGGCCACCAACCAAGAGGCTCTGTGGAAGTTCAGCAATCCTACTGTTCAAGGA 2499

QY 501 ACCTCGGACAAAGTCTAATCTGGAGCTGATCAGTCTAAGATGACCTGTGTGCTCGGACT 660
 DB 2500 ACCTCGGACAAAGTCTAATCTGGAGCTGATCAGTCTAAGATGACCTGTGTGCTCGGACT 2559
 QY 551 CTCCTCTGSCCTCTAATTAACCTCTTATCCGAAAAATGAAAAGGCTCTTCTTCAATA 720
 DB 2550 CTCCTCTGSCCTCTAATTAACCTCTTATCCGAAAAATGAAAAGGCTCTTCTTCAATA 2619
 QY 721 AAGNCTGACTACTATCAATTAATTAATGAGCCAGATGAAGTTCCTTTGATGAGCAGTGT 780
 DB 2620 AAGNCTGACTACTATCAATTAATTAATGAGCCAGATGAAGTTCCTTTGATGAGCAGTGT 2679
 QY 781 GAGCGCTCTCCCTTATGATGACCAAGTGGAGTTTCCCGGAGAGACTTAAAGTGGC 840
 DB 2680 GAGCGCTCTCCCTTATGATGACCAAGTGGAGTTTCCCGGAGAGACTTAAAGTGGC 2739
 QY 841 AATACACTTGGAGAGGGGCTTTTGGAAAAGTGTCAAGCATCAGCATTTGGCATTAA 900
 DB 2740 AATACACTTGGAGAGGGGCTTTTGGAAAAGTGTCAAGCATCAGCATTTGGCATTAA 2799
 QY 901 AATACACTTGGAGAGGGGCTTTTGGAAAAGTGTCAAGCATCAGCATTTGGCATTAA 960
 DB 2800 AATACACTTGGAGAGGGGCTTTTGGAAAAGTGTCAAGCATCAGCATTTGGCATTAA 2859
 QY 961 GAGTCAAGGCTCTGATGAGCTTAAATATCTTTGACCCACATTTGCCACCATCTGAAC 1020
 DB 2860 GAGTCAAGGCTCTGATGAGCTTAAATATCTTTGACCCACATTTGCCACCATCTGAAC 2919
 QY 1021 GTGTTTAACTCTGGAGGCTGCACCAAGCAGGAGGGCTCTGATGATGTTGTA 1080
 DB 2920 GTGTTTAACTCTGGAGGCTGCACCAAGCAGGAGGGCTCTGATGATGTTGTA 2979
 QY 1081 TACTGCAATATGAAAATCTCTCCAACTACCTCAAGAGCAAAAGTGCATTTATTTCTC 1140
 DB 2980 TACTGCAATATGAAAATCTCTCCAACTACCTCAAGAGCAAAAGTGCATTTATTTCTC 3039
 QY 1141 AACAGGATGACGACCTACACATGAGGCTTAAGAAAGAAATGAGCCAGGCTGGA 1200
 DB 3040 AACAGGATGACGACCTACACATGAGGCTTAAGAAAGAAATGAGCCAGGCTGGA 3099
 QY 1201 CAAGCAAGAAACCAAGACTAGATAGCTCACCAGCAGCAAGGCTTTGCCAGCTCCGC 1260
 DB 3100 CAAGCAAGAAACCAAGACTAGATAGCTCACCAGCAGCAAGGCTTTGCCAGCTCCGC 3159
 QY 1261 TTTACGAGGAATTAAGTCTGAGTGTGATGTTGAGGAGGAGGAGTCTGAGGGTTCTAC 1320
 DB 3160 TTTACGAGGAATTAAGTCTGAGTGTGATGTTGAGGAGGAGGAGTCTGAGGGTTCTAC 3219
 QY 1321 AAGGAGCCCATCAGTATGAGGAGTCTGATTTCTTACAGTTTTCAAGTGGCCAGAGCATG 1380
 DB 3220 AAGGAGCCCATCAGTATGAGGAGTCTGATTTCTTACAGTTTTCAAGTGGCCAGAGCATG 3279
 QY 1381 GAGTTCCTGCTCTCCAGAAAGTGCATTCAGGACCTGCGAGGAGAGAAATCTTTTA 1440
 DB 3280 GAGTTCCTGCTCTCCAGAAAGTGCATTCAGGACCTGCGAGGAGAGAAATCTTTTA 3339
 QY 1441 TCTGAGACACAGTGTGTGATTTGTGATTTGGCTTCCCGGAGATTTTAAAGAAC 1500
 DB 3340 TCTGAGACACAGTGTGTGATTTGTGATTTGGCTTCCCGGAGATTTTAAAGAAC 3399
 QY 1501 CCCGATTAATGTGAGAAAGAGATCTCGACTCTCTGAAATGGATGGCTCCGCAATCT 1560
 DB 3400 CCCGATTAATGTGAGAAAGAGATCTCGACTCTCTGAAATGGATGGCTCCGCAATCT 3459
 QY 1561 ATCTTTGACAAATCTACAGCACCAGGAGCGTGTGTTTACGGAGTATTTGCTGTGG 1620
 DB 3460 ATCTTTGACAAATCTACAGCACCAGGAGCGTGTGTTTACGGAGTATTTGCTGTGG 3519
 QY 1621 GAAATCTCTCTTAGTGGTCTCCATACCCAGGATCAAAATGATGAGACTTTTTC 1680
 DB 3520 GAAATCTCTCTTAGTGGTCTCCATACCCAGGATCAAAATGATGAGACTTTTTC 3579

Antagonists of KDR useful for treating diseases involving					
neovogenesis e.g. diabetic retinal vascularization, cancer					
(e.g. brain, breast, etc.) and forms of inflammation e.g. rheumatoid					
arthritis, psoriasis, contact dermatitis and hypersensitivity					
reactions. The polynucleotides are useful to screen for KDR					
antagonists/agonists and for gene therapy (e.g. by introducing a					
gene portion encoding a KDR protein containing functional ligand					
binding and membrane anchoring moieties but not tyrosine kinase					
activity). They are also useful to measure levels of human KDR.					
xx					
SQ	Sequence 4071 BP: 1169 A; 894 C; 1027 G; 981 T; 0 other:				
	Best Match	34.1%;	Score 715.4;	DB 20;	Length 4071;
	Query Local Similarity	65.5%;	Pred. No. 8e-207;		
	Matches 1078;	Conservative	0;	Mismatches 561;	Indels 6;
	Gaps				
OY	218	TCATGAATGTTCCCTCGACAGATTGAGGCACCTTATGCCCTGGAGAGCAGCAAGAIGTATACA	277		
Db	1886	TTAAGAATCATCTCTCGAGCAACGAGACTATGTCTGCTTGCTCAAGACAGGAGA	194		
I	278	CAGGGAAGAAATCCTCCAGAGAAGAAATACATCAGATCAGGAAGACACCATACC	337		
Db	1946	CCAGAAAAGACATGGCGTGGTCAGGCAGCTCACAGTCTTAGACGCTGTGGCCACCACGA	2905		
OY	338	TCCTCGAAACCTCAGTGTATCACAGATGGCCATCAGCAGITCCAACCTTTAGACTGTC	397		
Db	2006	TCACAGAACTCGAANTACACACACACAGATATGGGAAAGCATCGAAGTCTCATGCA	2665		
OY	398	ATGCTAATGGTGTCCCAGGCTCTAGATCACTTGGTTTTAAAAACAGCACAAAATACAAAC	457		
Db	2066	CGGCTCTGGGAATCCCGCTCCACAGATCATGTGTTTTAAGATAATGAGACCTTGATG	2125		
OY	458	AAGAGCTCGGAATATTTTAGCAGCAGAACGACGACGCTGTTATTAAGACAGTACAG	517		
Db	2126	AACACTCAGGCAITGTAATTGAGATGGAAACCGAACTCATCTCCGACAGTGGAGA	2185		
OY	518	AAGAGTGAAGGTGTCTATCACTCAAAGCCACCAACAGANGGCTCTGTGGAAGAT	577		
Db	2186	AGGAGACCAAGGCTCTACAGCTGCCAGCATGCGAGTGTCTTGCTGTGCAAAATGG	2245		
OY	578	CAGCATCCTCAGTGTTCAGGAACCTCGGCACAGTCTAATCTGGAGCTCATCTCTAA	637		
Db	2246	AGGCATTTTTCATATAGANGTGCACGAAAGACNAGCTTGGAAATCTATTTCING	2305		
OY	638	CATGCACTGTGTGGCTGGCACTCTCTCTGGCTCTATTAACCTCTTATCCGAAAAA	697		
Db	2306	TAGGCAAGCGGTCATGGCATGTCTCTCTGGCTACTTGTGTGTCATCATCTCAAGCAGC	2365		
Y	698	TGAAAAGTGC---TCTCTGAAATAAAGACTGACTACTATCAATTAATGACCCAG	754		
Db	2366	TAAAGCGGCCAATCGAGGGGAACTTGAGACAGGTACTCTGCTCTGTGACCCAG	2425		
OY	755	ATGAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCCCTTATGATGCCACGAAGTGGAGT	814		
Db	2426	ATGAATCCCAITGGATGAACTATGAAAGCACTGCTTATGATGCCACAAATGGGAAI	2485		
OY	815	TTCGCGGGGAGAGACTTAACTGGCAATACCTITGGGAAGAGGGGCTTTTGGAAAGTGG	874		
Db	2486	TCCCCAGAGACCGCTGAGCTAGGTAGGCTCTTGGCGTGGTGGCTTTGGCCAACTGA	2545		
OY	875	TTCAAGCATCAGCAATTTGGCAATTAGAAATACACTACGTCGGGACTGTGGCTGTGAAA	934		
Db	2546	TTCAAGCAGATGCTTTGGAATTTGAAGACAGCAACTTGCAGGACAGTAGCAGTCAAA	2605		
OY	935	TGCTGAAGAGGGGGSCACGGCCACGAGTACAAAGCTCTCATGACTGAGCTTAAATATCI	994		
Db	2606	TGTTGAAGAGAGGAGCAACACAGTAGCACTGAGCTCTCATGTGTGAATCTAAGATCC	2665		
OY	995	TGACCCACAAATGGCCACCATCTCAAGCTTGTTAACTCTCTGGAGGCTGCACCAAGCAG	1054		
Db	2666	TCATTCATATGCTCACCAATCTCAATGTGGTCAACCTCTCAAGTCTGTATCCAAAGCAG	2725		
OY	1055	GAGGCGCTCTGATGGTATGTTTGAATACTGCAAAATATGGAATCTCTCCACTACCTCA	1114		

Db 3443 GTGAGAGCCACGTTTTCAGAGTGTGGAGCAATTCGGAAATCTCTGCAAGCTAATG 3502

QY 1835 TACACAGAGTGTAAAGACTACAT 1859

Db 3503 CTCAGCAGATGCAAGACTACAT 3527

RESULT 5

ABL91754

ID ABL91754 standard: DNA: 4071 BP.

XX ABL91754;

AC ABL91754;

XX 28-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 97.

XX Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;

XX Plasmidium; virus; viroid; cytokine; prion; antisense oligonucleotide;

XX cytostatic; virucide; protozoacide; antibacterial; ds.

XX Homo sapiens.

XX DE10100586-C1.

XX 11-APR-2002.

XX 09-JAN-2001; 2001DE-1000586.

XX 09-JAN-2001; 2001DE-1000586.

XX (RIBO-) RIBOPHARVA AG.

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX WPI; 2002-270454/32.

XX Inhibiting gene expression in cells, useful for e.g. treating tumors,

XX by introducing double-stranded complementary oligoRNA having unpaired

XX terminal bases

XX Claim 13; Page 71-72; 104pp; German.

XX The invention relates to a method for inhibiting expression of a target

XX gene (ABL91658-ABL91797) in a cell by introducing at least one

XX oligoribonucleotide that has a double-stranded structure consisting of at

XX most 49 sequential nucleotide pairs, with at least part of one strand

XX complementary with the target gene and has at least one end a

XX single-stranded segment of 1-4 nt. The method provides

XX oligoribonucleotides for antisense inhibition of gene expression useful

XX e.g. for treating tumours but the oligoribonucleotides may also be

XX directed against genes present in pathogens (e.g. plasmidium or

XX viruses/viroids, pathogenic on humans, animals or plants) or against

XX cytokine, id, developmental or prion genes. The method provides more

XX effective inhibition of gene expression than use of known

XX oligonucleotides, probably because the unpaired overhang increases

XX stability and thus intracellular concentration.

XX Sequence 4071 BP; 1169 A; 894 C; 1025 G; 983 T; 0 other;

Query Match 34.0%; Score 713.8; DB 24; Length 4071;

Best Local Similarity 55.5%; Pred. No. 2.4e-206;

Matches 1077; Conservative 0; Mismatches 562; Indels 5; Gaps 2;

QY 218 TCATGATGTTTCCCTGCAAGATTCAGGCACCTATGCTGAGAGCCAGGAATGTATACA 277

Db 1886 TTAAGATGATCTCTTCAGAGCAAGAGAGACTATGCTGCTTGTCTCAAGACAGGAAGA 1945

QY 278 CAGGGAGAGAAATCTCCAGAGAGAAATACAAATCAGATCAGAGACAGGACCAIRCC 337

Db 1946 CCAGAAAGACATTCCTGGTGTAGGAGCTACAGCTCTTGTAGCGGTGTGACCCACCA 2005

QY 338 TCCTGCGAAACCTCAGTGTATCACACAGTGGCCATCAGCAGTTCACCACCTTATAGACIGTC 397

Db 2006 ICACAGAAACCTTGAGAAATCAGAGCAAGTATTCGGGAAGCATCGAGTCTCATGA 2065

QY 398 ATGCTAATGTTGTCGCGAGCTCAGATCACTTGGTTTAAAAACACCAAAATACAAAC 457

Db 2066 CGCATCTGGGAATCCCTCCACAGATCATGTGGTTTAAAGATAATGAGACCTTGTAG 2125

QY 458 AAGAGCTGSAATATTTTAGGACCAAGACAGCAGCTGTTTATTCAAGAGATCACAG 517

Db 2126 AAGACTCAGCAATGTATGAGAGTGGAAACCGGAACCTTACTATCCGAGAGTGA 2185

QY 518 AAGAGATGAGGTGTCTATCACTGCAAAAGCCACCAAGAGGCTCTGTGAAAGTT 577

Db 2186 AGGAGGAGGAGGCTCTACACCTGCCAGCATGCAGTGTCTTGGCTGTGCAAGAGTGS 2245

QY 578 CAGCATACCTCACTGTTCAGGAACCTCGACAGATCTTAATCTGGAGCTGATCACTCTAA 637

Db 2246 AGCATTTTTCATAATAGAGTCCCGAGAAAGAGCAAGTGTGAAATCATTTATCTAG 2305

QY 638 CATGCCACTGTGTGGCTCGGACTCTCTTCTGGCTCTCTTAACCTCTCTTATCCGAAAA 697

Db 2306 TAGCAGGCGGTGATGCGCATGTCTTCTGGCTACTTCTGTCACTCTAGGACCG 2365

QY 698 TGAAGAGTCTCTTCTTCTGAAATAAGACTGACTACCTATCAATTAATATGACCCAG 754

Db 2366 TTAGCGGCGCAATGAGGAGGAACTGAGACAGGCTACTTGTCCATCTCATGATCCAG 2425

QY 755 ATGAAGTCTCTTGTGAGTGTGAGCGCTCCCTTATGATGCCAGCAAGTGGAGT 814

Db 2426 ATGAATCTCTTGTGATGAACTGTGAAGCAGCTGCTTATGATGCCAGCAAGTGGAA 2485

QY 815 TTGCGCGGAGAGACTTAACTGGGCGAAATCACTTGAAGAGGGGCTTTTGGAAAAAGTGG 874

Db 2486 TCCACAGAGACCGGTGAAGCTAGGTAAAGCTCTTGGCGGTGGTCTTGGCCAGAGTGA 2545

QY 875 TTCAAGCATCAGCAATTCGCAATTAAGAAATCACTACCTAGCTGCGGAGTGTGGTGTGAAA 934

Db 2546 TTGAAGCAATGCTTTTGAATTCAGCAAGACAGCAACTTTCAGGACAGTATGAGTCAAAA 2605

QY 935 TGTGAAAGAGGGGCGCACGGCCAGGAGTACAAAGCTCTGATGCTGAGCTAAATCT 994

Db 2606 TGTGAAGAGAGGACACACAGTACAGTCTGAGTCTCATGTCTGAATCAAGATCC 2665

QY 995 TGACCCAAATTCGCGACCATCTGAACCTGTTAACTCTGCGAGCTCGCAACAGCAAG 1054

Db 2666 TCAATTCATATGCTCACCCTCAATGCTGCACTTCTAGTGGCTGTACCAAGCCAG 2725

QY 1055 GAGGCGCTCTGATGATGATGTTGAATCTGCAATATGGAATCTCTCCAACTACCTCA 1114

Db 2726 GAGGCGCTCTGATGATGTTGAATCTGCAATATGGAATCTCTCCAACTACCTCA 2785

QY 1115 AGAGCAAGCTGACTTATTTTTCTCAAGAGTGCAGCTACAGATGGAGCTTAAGA 1174

Db 2786 GGAGCAAGAGAAATCAATTTGTCCTCTACAAGACCAAGGGGCGAGTTCGTCAGAGGA 2845

QY 1175 AAGAAAAATGAGCGAGCGCTGCAAGAGGCAAGCAAGCAAGCAAGTATAGTGGTCAACA 1234

Db 2846 AAGACTACCTTGGAGCAATCCCTG---TGGATCTGAAGCGCGCTTGGAGCATCACTCA 2902

QY 1235 GCAGCAAGAGCTTTTCGAGCTCCGGCTTTCAGCAAGATAAAAGTCTGAGTGTGGG 1294

Db 2903 GTAGCCAGAGCTCAGCCAGCTCTGGATTTGTGGAGAGAGTCCCTCAGTGTAGTAGAG 2962

QY 1295 AAGAGAGGATTTCTGACGGTTTCTACAAGAGCCCATCACTATGGAAGATCTGATTTCT 1354

Db 2963 AAGAGGAAGCTCTCAAGATCTGTATAAGGACTTCTGACCTTGGAGCATCTCATCTGTT 3022

QY 1355 ACAGTTTCAAGTGGCGAGCGCATGGATTCCTGCTTCTCCAGAAAGTGCATTCATCCGG 1414

Db 3023 ACAGTTTCAAGTGGCGAGCGCATGGATTCCTGCTTCTCCAGAAAGTGCATTCATCCAG 3082

QY 1415 ACCTGCGAGAGAAACATCTTTTATCTGAGAAACAGTGTGAGATTTGTGATTTG 1474

Db 2306 TAGGCACGACGGTGAATGCCATGTTCTTGGCTACTTCTGTCAATCACTCTAGGACCG 2365
Qy 698 TGAAGAGTC---TTCTTTGTAATAAGACTGACTACTATCAATATAATAGGAGTAC 754
Db 2366 TTAACGGGGCAAATGGAGGGGAATGAAGACAGGCTACTTGTCCATGCTGATGAC 2425
Qy 755 ATGAAGTTCCTTTGGAAGAGCAGTGTGAGCGGCTCCCTTAATGATGCCAGAGTGGAGT 834
Db 2426 ATGAAGTTCCTTTGGAAGAGCAGTGTGAGCGGCTCCCTTAATGATGCCAGAGTGGAGT 845
Qy 815 TTGCCGGGAGAGCTTTAAATCGGCAATACATTTGAAGAGGGGCTTTTGAAGAAGTGG 874
Db 2486 TCCCGAGAGAGCGGCTGAGCTTAGTACCTCTTGGCCGTGTGCTTTCGCCAAGAGA 2545
Qy 875 TTAAGATCAGATTTGGCAATTAAGAAATACCTAGTGTGCGGAGTGTGCTGTGAAGA 934
Db 2546 TTGAAGCAGATGCTTTTGAATTAAGAAATACCTAGTGTGCGGAGTGTGCTGTGAAGA 2605
Qy 935 TCTGAAGAGGGGCGCCAGCGGCTGAGTACAGCTCTGATGAGTGAATTAATCT 994
Db 2606 TGTGAAGAGGAGCAGCAGAGTGTGAGCTGTGAGCTGTGAGTGTGAGTGTGAGTGTG 2665
Qy 995 TGACCCCAATTTGGCAGCAGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 1054
Db 2666 TCATTCATATTTGGTCAACCTCTCAATGTGCTGAGCTGTGAGCTGTGAGCTGTGAG 2725
Qy 1055 GAGGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1114
Db 2726 GAGGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2785
Qy 1115 AGAGCAACGCTGACTTTATTTTCTCAAGAGGATGAGCTACACATGAGGCTTACGA 1174
Db 2786 GGAGCAGAGAAATGAATTTTCTCAAGAGGATGAGCTACACATGAGGCTTACGA 2845
Qy 1175 AGAGCAACGCTGACTTTATTTTCTCAAGAGGATGAGCTACACATGAGGCTTACGA 1234
Db 2846 AAGACTAGCTTTGGAGCAATCCCTG---TGAATCTGAAGAGGCTTGGAGCAATCCT 2902
Qy 1235 GCAGCGAAGCTTTGGAGCAATCCCTG---TGAATCTGAAGAGGCTTGGAGCAATCCT 1294
Db 2903 GTAGCCAGCTTACAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2962
Qy 1295 AAGAGAGGATTTGAGGCTTTTCTCAAGAGGATGAGCTACACATGAGGCTTACGA 1354
Db 2963 AAGAGAGGATTTGAGGCTTTTCTCAAGAGGATGAGCTACACATGAGGCTTACGA 3022
Qy 1355 ACAGTTCCTGAGCTTACAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1414
Db 3023 ACAGTTCCTGAGCTTACAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3082
Qy 1415 ACCTGGCAGGAGAAATCTTTTCTGAGCAACAGCTGTGAGTGTGAGTGTGAGTGTGAG 1474
Db 3083 ACCTGGCAGGAGAAATCTTTTCTGAGCAACAGCTGTGAGTGTGAGTGTGAGTGTGAG 3142
Qy 1475 GCTTGGCAGGAGAAATCTTTTCTGAGCAACAGCTGTGAGTGTGAGTGTGAGTGTGAG 1534
Db 3143 GCTTGGCAGGAGAAATCTTTTCTGAGCAACAGCTGTGAGTGTGAGTGTGAGTGTGAG 3202
Qy 1535 CTCTGAATGATGCTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1594
Db 3203 CTCTGAATGATGCTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3262
Qy 1595 TGTGCTTACGAGTGTGCTTGGGAAATCTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTG 1654
Db 3263 TGTGCTTACGAGTGTGCTTGGGAAATCTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTG 3322
Qy 1655 GAGTCAAAATGATGAGTGTGCTTGGGAAATCTTCTTCTGAGTGTGAGTGTGAGTGTGAG 1714
Db 3323 GAGTCAAAATGATGAGTGTGCTTGGGAAATCTTCTTCTGAGTGTGAGTGTGAGTGTGAG 3382
Qy 1715 CTGAGTACTTACTCTTGAATCTATCAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1774

Db 3383 CTGATTATATCTACACAGAAATGTACACAGACCATCTGCGACTCTGCGACGGGAGCCCA 3442
Qy 1775 AAGAAAGCCCAAGATTTGCAAGACTTTGCAAAACTAGTGTGATTTGCTTCAAGCAATG 1834
Db 3443 GTCAGAGACCCACGCTTTTCAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3502
Qy 1835 TACAACAGGATGTTAAAGACTACAT 1859
Db 3503 CTCAGCAGGATGTTAAAGACTACAT 3527
RESULTS
AAV34763
ID AAV34763 standard; DNA: 4236 BP.
XX AAV34763;
XX AC
XX 27-AUG-1998 (first entry)
XX DE Human KDR genomic DNA.
XX Kinase insert domain containing receptor; KDR: screening; inhibitor;
KW Vascular endothelial growth factor; VEGF; angiogenesis; treatment;
KW cancer; ss.
XX OS Homo sapiens.
XX FH Key
XX FI CDS 1..4071
XX FT /+tag- a
XX FJ /product= KDR
XX FT /note= "kinase insert domain containing receptor"
XX PN US5766860-A.
XX PD 16-JUN-1998.
XX PF 25-FEB-1997; 97US-0810116.
XX PR 23-NOV-1992; 92US-0930548.
XX PR 25-FEB-1997; 97US-0810116.
XX PA (AMCY) AMERICAN CYANAMID CO.
XX PI Carillon ME, Terman BI;
XX DR WPI: 1998-361682/31.
XX DR F-PSDB; AAW59275.
XX PT Screening assay for vascular endothelial cell growth factor
XX PI antagonists - using recombinant cells expressing receptor protein
XX PS Claim 1: Fig 7A-M: 51pp; English.
XX CC This sequence encodes a novel human growth factor receptor, kinase
CC insert domain containing receptor or KDR. This receptor is capable of
CC binding to the vascular endothelial cell growth factor, VEGF and is
CC used in a screening assay which identifies compounds that inhibit VEGF
CC action on KDR. Such compounds which inhibit binding of VEGF to the KDR
CC may inhibit angiogenesis and thus be useful for treating cancer.
XX SQ Sequence 4236 BP: 1215 A: 938 C: 1062 G: 1020 T: 0 other;
Query: Match 33.8%; Score 710.6; DB 19; Length 4236;
Best Local Similarity 65.3%; Pred No. 2.4e-205;
Matches 1075; Conservative 0; Mismatches 564; Indels 6; Caps 2;
Qy 218 TCATGAATGTTTCTCTCCAGATTCAGGCACTATGCTTCAGAGCCAGCAATGTATACA 277
Db 1896 TTAAGAAATGCAATCTTCCAGGACCAAGGAGACTATGCTGCTTCTCAAGACAGGAAGA 1945
Qy 278 CAGCGGCAAGAAATCTCTCCAGAAAGAAATTAACAATCAGATCAGGAGCCATACC 337

Db 1945 CCAGAAAGACATTCGGTGTGACGAGCTACAGTCTCAGAGTGTGGCACCACGA 2005
QY 338 TCTTCCGAAACCTCAGTGTATCAGCAGTGGCCATCAGCAGTTCACACCTTACGACTTC 397
Db 2005 TCACAGGAAACCTGGAGATCAGACACAGTATGGGGAAGCATCGAGTCTCATGCA 2065
QY 398 ATGCTATGTTGTCGCGGCTCAGTACCTTGGTTTAAACACACCAAAATACAC 457
Db 2065 CGCATGTGGAAATCCCTCCACAGATCATGTGTTTAAAGATATGAGACCTTGTAG 2125
QY 458 AAGACCTTGGAAATTTTAGCAGCAGGACGACGCTGTTTATGAAAGATCAG 517
Db 2125 AAGACTCAGCAATGTATGAGGATGGGACCGGACCTCCTATCCGAGGTGAGGA 2185
QY 518 AAGAGATGAGGTGCTTACTTCCAAAGCCCAACACAGAGGCTCTGTGGAAAT 577
Db 2196 AGCAGGAGGAGGCTCTACCTCCAGGATGCTGTGCTGTGCTGTGCAAAAGTG 2245
QY 578 CAGATACCTCAGCTGTTCAGGACCTCGGACAGCTTAATCTGAGGTGATCCTNA 637
Db 2245 AGCAATTTTCAATAGAGGTGCCAGGAAAGAGCAATTTGAAATCATATTTCTAG 2305
QY 638 CATGACCTGTGGTGGGACTCTCTCTGCTCTTATTAACCTCTTATCCGAAAAA 697
Db 2305 TAGGACGACGCTGATGCTGCTCTCTGCTCTTATTAACCTCTTATCCGACCG 2365
QY 698 TGAAGATC---TCTCTGAAATTAAGACTGACTACCTATCAATATATGAGCCAG 754
Db 2366 TTAAGCGGCAATGAGGAGGAACTGAAGACAGGCTACTTGTCCATGCTATGATCAG 2425
QY 755 ATGAAGTCTCTTGGATGAGGAGTGTGGGCTCCCTTAATGATGCGACAGTGGGAGT 814
Db 2425 ATGAATCTCCATGGATGAACATGTGAGCAGCTGCCCTTAATGATGCGACAAATGGGA 2485
QY 815 TTGCGCGGAGAGACTTAACTGGGCAATCATTGGAAGAGGGCTTTTGGAAAAAGTGS 874
Db 2486 TCCCGAGAGCGGCTGAACCTAGTAAAGCTCTTGGCGGTGGTCTTGGGCAAGAGA 2545
QY 875 TTCAGCATCAGCAATTTGGCATTAAGAAATCAGCTAGCTGCGGACTGTGGTGAAGA 934
Db 2546 TTGAAGCAGATGCTTTGGAATTTGCAAGACAGCACTGTGAGGACAGTACAGTCAAAA 2605
QY 935 TGCTGAAGAGGGGCGCAGCGGCTGAGTACAAAGCTCTGATGACTGAGTAAATTCI 994
Db 2605 TGTGGAAGAGGAGCAACACAGAGTACGAGCTCTGATGCTGAGTCAAGTCAAGTCC 2665
QY 995 TGACCCACATGCTGAGCATCTGAGCTGTTTACCTGTGGAGGCTTGCACCAAGCAAG 1054
Db 2665 TCATCATATTTGTCACCACTCAATGTTGTTCAACCTCTAGTGTGCTGACCAAGGCG 2725
QY 1055 GAGGCTCTGATGCTGATTTGATGTAATCTGCAAAATATGAAATCTCTCCACTACTCA 1114
Db 2726 GAGGCGCACTATGCTGATTTGTAATCTGCAAAATTTGAAACCTGTCCACTTACCTGA 2785
QY 1115 AGAGCAACGTGACTTATTTTCTACAGAGTACGACATACAGTACGAGTCAAGA 1174
Db 2786 GGAGCAAGAGAAATGAATTTGCTCCCTACAGAGCAAAAGGGCCACTTCCGTCAGGGA 2845
QY 1175 AGAGAAATCGACCGAGCTTGAAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1234
Db 2846 AAGACTGTTTGGAGCAATCCCTG---TGAATCTGAAGAGGGCTGTGAGAGTACCA 2902
QY 1235 CGAGCAAGGCTTTGAGCTCCGCTTTTACGAGGAAATGAAGTCTGATGATGTGAGG 1294
Db 2903 GTAGCCAGGCTCAGCCAGCTCTGGAATTTGTGAGGAGAGAGTCCCTCAGTGTAGAG 2962
QY 1295 AAGAGGAGATCTTGAGGTTTCTACAGAGGAGCCATCAGTATGAGAGATCTGATTT 1354
Db 2963 AAGAGAGAGCTCTGAGATCTGTAAGGACTTCTTCCAGCTTTGGAGCACTCTCATCT 3022
QY 1355 ACAGTTTTCAGTGGCCAGAGGCTGAGTCTCTGCTTCCAGAAAGTCCATTCATCGG 1414
Db 3023 ACAGTTTTCAGTGGCTTGGAGGCTAGGAGTCTTGGATCGGAAAGTATCCACAGG 3082

QY 1415 ACTGCGAGCGAGAAACATCTTTATCTGAGAAACAGCTGTGTGAGATTTGTGTTTG 1474
Db 3083 ACTGCGCGGACGAAATATCTCTTATCGGAGAAACAGCTGGTAAATCTGTGACTTG 3142
QY 1475 GCTTGGCCGCGGATATTTATGAACCCCGATTTATGTGAGAAAAGAGATATCGACTTC 1534
Db 3143 GCTTGGCCGCGGATATTTATGAATCCAGATTTATGTGAGAAAAGAGATGCTCGCTCC 3202
QY 1535 CTCTGAATGATGCTCCCGATCTATCTTTGACAAATCTACAGCACCAGAGGACG 1594
Db 3203 CTITGAATGATGCGCCAGAACAAATTTTIGACAGATGTACAAATCCAGAGTGACG 3262
QY 1595 TGTGCTTACGAGTATTTCTGTGGAAATCTTCTCTTGTGTTGTTCTTCAATCCAG 1654
Db 3263 TCTGCTCTTTTGGTGTCTGTGGGAAATATTTCTTGTGTTGTTCTTCAATCCG 3322
QY 1655 GAGTACAAATGATGAGGACTTTTCAGTCCGCTGAGGAGGAGCATGAGGATGAGCTC 1714
Db 3323 GGTAAAGATTTGATGAAGATTTTGTAGCGATTTGAAGAAGGAAGTATGATGAGGCGCC 3382
QY 1715 CTGAGTACTCTCTCTGAAATCTATCAGATCATCTGAGCTCTGCGACAGACCCAA 1774
Db 3383 CTGATTTACTACACCAAGATGTACAGACCATCTGAGCTCTGCGACGCGGAGCCAA 3442
QY 1775 AAGAAAGCCGAGATTTGAGCAACTTTGGAAGAACTAGTGTATTTCTTCAAGCAATG 1834
Db 3443 GTGAGAGCCGACCTTTTTCAGAGTTGGTGAACATTTTGGAAATCTCTTGAAGCTAATG 3502
QY 1835 TACAACAGGATGTAAGACTACAT 1859
Db 3503 CTCAGCAGATGGCAAGACTACAT 3527

RESULT 9

AA138735
ID AA138735 standard; cDNA; 5404 BP.

XX AC AA138735;

XX DT 11-DEC-1996 (first entry)

XX DE Murine foetal liver kinase 1 cDNA.

XX KW Murine: foetal liver kinase 1; fik-1; protein tyrosine kinase;
XX KW monoclonal; antibody; extracellular domain; receptor assay;
XX KW haematopoietic stem cell; ligand; stimulation; proliferation;
XX KW differentiation; treatment; anaemia; bone marrow damage;
XX KW cancer chemotherapy; radiation; ds.

XX CS Mus musculus.

XX FH Key Location/Qualifiers

FI CDS 208..4311

FI sig_peptide 208..264

FI sig_peptide 265..4308

FI sig_peptide 265..4308

XX FN 15548065-A.

XX FD 20-AUG-1996.

XX PF 02-APR-1991; 91US-0679666.

XX PF 19-NOV-1992; 92US-0977451.

XX PF 02-APR-1991; 91US-0679666.

XX PF 28-JUN-1991; 91US-0728913.

XX PF 15-NOV-1991; 91US-0793065.

XX PF 24-DEC-1991; 91US-0813593.

XX PF 26-JUN-1992; 92US-0906397.

XX PF 12-NOV-1992; 92US-0975049.

PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252517.
 XX (UYP-R) UNIV PRINCETON.
 XX Lemischka IR;
 XX WPI; 1996-392678/39.
 DR P-PSDB; AAT38735.
 XX Anti-fetal liver kinase 2 (flk-2) antibodies - useful in assays,
 PT for isolating haematopoietic stem cells expressing receptor and for
 PT obtaining ligands
 XX Disclosure: Columns 51-62; 50pp; English.
 XX The present sequence encodes murine foetal liver kinase 1 (flk-1),
 CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
 CC raised against the extracellular portion of flk-1 can be used to
 CC assay for flk receptors on the surface of haematopoietic stem
 CC cells, and to isolate positive cells. The antibodies can also
 CC be used as, or to obtain ligands, which stimulate the proliferation
 CC and/or differentiation of stem cells. The ligands can be used, e.g.
 CC for treating anaemia, or bone marrow damage resulting from cancer
 CC chemotherapy, or radiation.
 XX Sequence 5404 BP: 1411 A; 1297 C; 1423 G; 1273 T; 0 other:
 SQ
 Query Match 33.7%; Score 708.4; DB 17; Length 5404;
 Best Local Similarity 65.0%; Pred. No. 1.3e-204;
 Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;
 QY 201 CACTTATATCTTACCATCATGATGTTTCCCTGCAAGATTCAGGCACCTATGCGTGCAG 260
 DB 2070 CATCTGATTTGGCAITTCAGATGCTCTCTGAGGACCAAGGCGACTATGTTGCTC 2129
 QY 261 AGCCAGGAATGTATACACAGGCGGAAGAAATCTCCAGAGAAAGAAATTAACAATCAGAGA 320
 DB 2130 TGCTCAAGATAGAGACCAAGAAAGACATTCCTGGTCAACAGCTCAITCACTCTAGA 2183
 QY 321 TCAGGAGACCAATACCTCTCGGAACTCAGTATCACACAGTGGCCATCAGAGTTC 380
 DB 2190 GCGCATGGCACCCATGTATCACCGGGAATCTGGAGAAATCAGACAACCAATGGCGAGAC 2249
 QY 381 CACCACCTTTAGACTGTCAIGCTTAATGTGTCCCGAGGCTCAGATCACTTGGTTAAAAA 440
 DB 2250 CATTGAAGTGACTTGGCCAGACATCTGMAATCTACCCCAACATATACATGGTTCANAGA 2309
 QY 441 CAACCAAAAATACAACAGAGCTGGAATATTTTAGGACCAAGAGCAGCAGCTGT 500
 DB 2310 CAACGAGACCCCTGGTAGAAGATTCAGGCATTTGACTGAGAGATGGGAACCGGAACCTGAC 2369
 QY 501 TATTGAAGAGTCAACAGAGAGATGAAGTGTCATCACTGCAAGGCCACCAACCAAGAA 560
 DB 2370 TATCCGAGGTTGAGGAGGAGATGGAGGCTCTACCTTGCCAGGCTTGCAATGCT 2429
 QY 561 GGCTCTGTGGAAGTTCAGATACCTCAGTCTTTCAGGAACTCCGACAAAGCTAAATCT 620
 DB 2430 TGGCTGTGCAAGCGGAGACGCTCTTCATATAGAGGTGCCAGGAAAGAACCAACTT 2489
 QY 621 GGAGCTGATCACTTACATGCACCTGTGGCTGGGACTCTCTCTGGCTGCTTATTAAC 680
 DB 2490 GGAAGTCAATCTCTGCGCACTGAGTATGCAATGCTCTGCTGCTGCTGCTGCT 2549
 QY 681 CCTCTTATCCGAAAAATGAAGAGTC- - -TCTCTGAAATGAAGACTGACTACCTATC 737
 DB 2550 CATCTCTGACGAGCCGTTTACGGGGCCAAAGAGGGCACTGAGACAGGCTACTTGTC 2609
 QY 738 AATTAATATGGACCCAGATGAAGTCTCTTTGATGAGCAGTGTGAGCGGCTCTTATGA 797
 DB 2610 TATGTCAATGGATCCAGATGAATGGCCCTGGATGAGCGCTGTGAGCGCTGCTATGA 2663
 QY 798 TGGCAGCAAGTGGGAGTTTGCCCGGAGAGACTTAACTGGGCATAATCATCTGGAGAG 857

DB 2570 TCCAGCAAGTGGGAATTCGCCAGGACCGGCTGAACCTAGGAAAAACCTCTTTGGCGCGG 2729
 QY 858 GGCTTTTGAAAAAGTGTTCAGACATCAGCATTTGGCAATTAAGAAATCACCCTACCTGCGG 917
 DB 2730 TGGCTTCGGCAAGTATTAGCGCACACGCTTTTGGAAATTCACAACAGACGACTTGCAA 2789
 QY 918 GACTGTGGCTGTGAAATGCTGAAAGAGGGGCCACGGCCACGAGGTACAAAGCTCTGAT 977
 DB 2790 AACAGTAGCGGTCAAGATGTTGAAAGAGGAGCAACACACAGCGAGCATCGAGCCTCAT 2849
 QY 978 GACTGAGCTTAAATCTTGACCCACATTCGCCACCATCTGACCTGGTAACTCTCTGG 1037
 DB 2850 GTCTGAATCTAAGATCTCTCAACATTTGTCACCATCTCAATGTGGTGAACCTCTCAGG 2909
 QY 1038 AGCTGCACCAAGAGGAGGCTCTGATGGTGAATTTGTAATCTGCAATATGGAAT 1097
 DB 2910 CGCTGCACCAAGCGGAGGGCTCTCATGTGATTTGGAATTTCTGGAATTTGGAAA 2969
 QY 1098 TCTCTCAACTACCTCAAGAGCAACGTGACTATTTTCTCAACAGGATGCAGACT 1157
 DB 2970 CCTATCAACTTACTTACGGGCAAGAGAAATGAATTTGTTCCCTATAAGAGCAAAAGGGC 3029
 QY 1158 ACATGAGGAGCTTAAGAAAGAAATGGAGCGAGGCTGGAACAGGCAAGAAACCAAG 1217
 DB 3030 ACGTCTCCGCCAGGCAAGACTACGTTGGG- -GAGCTCTCCGTGATCTGAAAGAGC 3086
 QY 1218 ACTAGTATGCTCACAGCAGGCAAGACTTTGCGAGCTCCGGCTTTTCAGGAGATAAAG 1277
 DB 3087 CTITGGAGCATCTACCCAGCAGCAGAGCTCTGCCAGCTCAGGCTTTGTGAGGAGAAATC 3146
 QY 1278 TCTGAGTGAATGTGAGGAGAGGAGATCTCAGGCTTTCTACAGGAGGCCATCACTAT 1337
 DB 3147 GCTCAGTGAATGTAGGAGAGAAAGCTTCTGAGAACTGTACAGGAGCTTCTGACCTT 3206
 QY 1338 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGGAGCAATGAGTTCCTCTCTCCAG 1397
 DB 3207 GGAGCACTCTCTGTACAGCTTCCAAGTGGCTAAGGCAATGGAGTCTTGGCATCAAG 3266
 QY 1398 AAGTGCATTCATCGGACCTGCGACGCGAGAAACATCTTTATCTGAGAACAACTGGT 1457
 DB 3267 GAAGTGTATCCAGAGGAGCTTGGCAGCAGCAACATCTCTCTATCGGAGAAAGTGTGT 3326
 QY 1458 GAAGATTTGTGATTTTGGCTTTCGCCGGATATTTTAAAGAACCCGATTTATGTGAGAAA 1517
 DB 3327 TAAGATCTGTGACTTCGCTTTGGCCGGGACATTTTAAAGACCCGGATTTGTGAGAAA 3386
 QY 1518 AGGATATCTCGACTTCTCTGAAATGGATGGCTCCCGAATCTATCTTTGACAAAATCTA 1577
 DB 3387 AGGATATCGCCACTTCCCTTTGAAGTGGATGGCCCGGAAACCAATTTTTCACAGAGTATA 3446
 QY 1578 CAGCAACCAAGAGCAGCTGTGGTCTTACGGAGTATTGCTGTGGGAAATCTTCTCTCTAGG 1637
 DB 3447 CACAATTCAGACGGAATGTGGTCTTTCGGTGTGTGCTCTGGGAAATATTTTCTTAGG 3506
 QY 1638 TGGGTCTCCATACCAGAGGTACAAATGGATGAGGACTTTTGCAGTCCGCTGAGGAGAG 1697
 DB 3507 TGGTCTCCCATACCCTGGGGTCAAGATTTGATGAAGAAATTTTGTAGGAGATTGAAGAAGG 3566
 QY 1698 CATGAGGATCAGAGCTCTGAGTACTCTACTCTGAAATCTATCAGATCATCTGAGACTG 1757
 DB 3567 AACTAGATTCGGGGCTCTGACTACTACCTCCAGAAATGTACCAGACCAATGCTTGAGCTG 3626
 QY 1758 CTGSCACAGAGACCCCAAGAAAGAGCCAGATTTGCAAGACTTGTGGAAAAAATAGGTGA 1817
 DB 3627 CTGGATGAGGACCCCAACAGAGACCCCTCGTTTTCAGAGTTGGTGGAGCATTTGGAAA 3686
 QY 1818 TTGCTTCAAGCAATGTACAAACAGGATGGTAAGAGCTACAT 1859
 DB 3687 CCTCTTGAACCAATTCGCACAGGATGCAAGAGTATATAT 3728

RESULT 10

AA77516
ID AAX77516 standard; cDNA; 5404 BP.

AC AAX77516;

DT 05-AUG-1999 (first entry)

DE Murine flk-1 cDNA.

XX Murine: flk-2; flk-1; cell isolation; fetal liver kinase: receptor;
KW monoclonal; polyclonal; antibody; tyrosine kinase; ds.

OS Mus sp.

XX Key Location/Qualifiers
FT 208...4311
FT /*tag= a
FT /product= "flk-1"

XX US5912133-A.

XX 15-JUN-1999.

XX 10-FEB-1998; 98US-0021324.

XX 19-NOV-1992; 92US-0977451.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 26-JUN-1992; 92US-0906397.

XX 12-NOV-1992; 92US-0975049.

XX 30-APR-1993; 93US-0052669.

XX 31-OCT-1994; 94US-0252498.

XX 15-FEB-1996; 96US-0601891.

XX (UYPF-) UNIV PRINCETON.

XX Lemishka IR:

XX WPI; 1999-357184/30.

XX P-PSDB; AAY08618.

XX Isolating hematopoietic cells expressing fetal liver kinase 1

XX receptors
XX Disclosure: Fig 2; 59pp; English.

XX This invention describes a novel method of isolating cells expressing

XX fetal liver kinase 1 (flk-1) receptors on their surface and comprises

XX binding the cells to a polyclonal or monoclonal antibody specific to

XX the flk-1 receptor and isolating the cells that have bound to the

XX antibody. The method can be used to isolate hematopoietic stem cells in

XX any mammal but preferably a rat, mouse, rabbit or human. The proteins of

Db 2190 GCGCATGGCACCACATGATCACCGAAATCTGGAGAATACAGAAACCAACCTTGGCGAGAC 2249
Qy 381 CACCACITTTAGACTGTCTGCTAATGGTGTCCCGAGCCTCAGATCACTTGTGTTTAAAAA 440
Db 2250 CATTGAAGTGACTTGGCCAGCACTGGAAATCTCTACCCACACATTTACATGTTCAAGA 2309
Qy 441 CAACCAAAATACACAGAGCCTGGAAATTTTATAGCAGGAGGAGCAGCAGCTGT 500
Db 2310 CAACGAGACCTTGGTGAAGATTTCAGGCAITGTACTGAGAGATGGAAACCGGAACCTGAC 2369
Qy 501 TATTGAAGAGTCAACAGAGAGGATGAAGGTGTCTATCACTGCAAGAGCCCAACACAGAA 560
Db 2370 TATCGCAGGAGTGAAGAGAGGATGGAGGCTCTACACCTGCCAGGCTGCAATGTCT 2429
Qy 561 GGGCTGTGTGGAAGTTTCAAGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 620
Db 2430 TGGGTGTGCAAGAGCGGAGACGCTCTTCAATAAGAGGTGCCAGGAAAGACCAACTT 2489
Qy 521 GGAGCTGATCACTCTAACATGCACTGTGTGGCTGGAGCTCTCTCTCTCTCTCTCTCTCT 680
Db 2490 GGAAGTCAITATCTCTGCGCACTGCAAGTGAATGCAATGTTCTCTCTCTCTCTCTCT 2549
Qy 581 CTTCTCTTATCGGAAATGAAAAGGTC---TCTCTCTGAAATAAAGACTGACTACTCTATC 737
Db 2550 CATTCTCGTACGAGCGCTTACGGGGCAATGAAGGGAACTGAACAGACAGCTACTTGT 2609
Qy 738 AATTAATAGCACCAGATGAAGTTCTCTTTGGATGAGCAGTGTGAGCGCTCTCTCTCTCT 797
Db 2610 TATTCGCAAGTCCAGATGAATGCTCTTGGATGAGCAGTGTGAGCGCTCTCTCTCTCT 2669
Qy 798 TCCAGCAAGTGGAGTTTCCCGGAGAGAGCTTAACTGGGCAATCACTTGGAAAGG 857
Db 2670 TCCAGCAAGTGGAGTTTCCCGGAGAGAGCTTAACTGGGCAATCACTTGGAAAGG 2729
Qy 858 GCGTTTGGAAAGTGTCAAGCATGACGATTTGGCAATTAAGAAATCACTTGTGCTGCG 917
Db 2730 TCCCTCGGCAAGTGAATGAGCAGAGCGCTTTTGAATGACAGACAGAGCTTGTCAA 2789
Qy 918 GACTGTGCTGTGAATGCTGAAGAGGGGGCCAGCGGAGAGTACAAGCTCTCAT 977
Db 2790 AACAGTACCGCTCAAGATGTTGAAGAGAGAGAGCAACACAGCAGCAGCTATCGAGCCTCAT 2849
Qy 978 GACTGTGCTGTGAATGCTGAAGAGGGGGCCAGCGGAGAGTACAAGCTCTCAT 1037
Db 2850 GTCTGAGTCAAGATCTCATCCATTTGGTCACTCTCATGTTGGTGAACCTCTCTAGG 2909
Qy 1038 AGCTGTGCAAGCAAGAGAGGCGCTCTGATGCTGATTTGTAATCTGCAATATGGA 1097
Db 2910 CGCTGTGCAAGCAAGAGAGGCGCTCTGATGCTGATTTGTAATCTGCAATATGGA 2969
Qy 1098 TCTGTGCAAGTCAAGAGCAAGAGTGAATTTTCTCAACAGAGATGAGCAGCT 1157
Db 2970 CCTATCAACTTACTTACGGGCAAGAGAAATGAATTTGTTCCCTATAAGAGCAAGGCGC 3029
Qy 1158 ACACATGAGAGCTTGAAGAAAGAAATGAGCGGCTGGAAACAGCAAGCAACCAAG 1217
Db 3030 AGCTGTGCAAGCAAGAGAGTACGTTGGG---GAGCTCTCGGTGATCTGAAAGAGC 3086
Qy 1218 ACTATGAGGTACACAGAGCAAGAGCTTTGCGAGCTCTCGGCTTTTCAGGAAGATAAAG 1277
Db 3087 CTTGACAGCAATCACCAGCAGCAGAGCTCTGCCAGCTCTGTTGTTGAGGAGAAATC 3146
Qy 1278 TCTGAGTGAAGTGAAGAGAGAGGATTTCTGAGGTTTCTCAAGGAGCCATCAT 1337
Db 3147 GCTCAGTGAAGTGAAGAGAGAGAGCTTTCTGAAGAACTGTACAGAGCTTCTGAGCTT 3206
Qy 1338 GGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1397
Db 3207 GGAGCACTCACTGTACAGCTTCCAGTGGCTGAAGGCAATGAGTCTTGGCATCAAG 3266
Qy 1398 AAGTGCATTCATCGGAGCTTGGCAGGAGAGAAATTTCTTTTATCTGAGAACAGCTGCT 1457
Db 1457

Query Match 33.7%; Score 708.4; DB 20; Length 5404;
Best Local Similarity 65.0%; Pred. No. 1.3e-204;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;

Qy 201 CACTCTTAATCTTACCAATCAATGATTTCTCTGCAAGATTCAGGCACTTGTCTGAG 260

Db 2070 CATCTGTGATGTGGCAATTCAGAAATGCTCTCTGAGGACCAAGCGGACTATGTTGCTC 2129

Qy 261 AGCCAGGAATGATACAGAGGGAAGAAATCCTCCAGAAAGAAATTAACATACAGAGA 320

Db 2130 TCTCAAGATGAAGAGACCAAGAAAGACATGCTGCTCAACAGCTCAATCACTTGA 2189

Qy 321 TCAGGAGCACCATTACCTCTCTGGAAGCTTCACTGATACACAGTGGCCATCAGCAGTTC 380

Db 2790 AACAGTAGCGGTCAAGATGTTGAAGAAGGAGCAACACACACACGAGCAICAGGCCICAT 2849
Qy 978 GAGTACGCTAAATAATCTTACCCACCATTCGTCACCATCTGACCTGTTAACTGCTGG 1037
Db 2850 GTCTGAATCAAGATCTCTTCAATCTTGGTCAATCTGACCTGTTAACTGCTGG 2309
Qy 1038 AGCTGCAACAGCAAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1097
Db 2910 CGCTGCAACAGCAAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2959
Qy 1038 TCTCTCACTACCTCAACAGCAAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTT 1157
Db 2970 CCTATCACTTCTTACGGGCAAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTT 3029
Qy 1158 ACATGAGGCTTAAAGAAAGAAATGAGGCGGCTGGAAGCAAGCAAGCAAGCAAG 1217
Db 3030 ACGTCTCCGCAAGGCAAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3086
Qy 1218 ACTAGATGCTGCAAGCAAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1277
Db 3087 CTGACAGCAATCAAGCAAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3146
Qy 1278 TCTGATGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1337
Db 3147 GCTCAGTCTGATGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3206
Qy 1338 GGAAGTCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1397
Db 3207 GGAAGTCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3266
Qy 1398 AAGTCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1457
Db 3267 GAGTCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3326
Qy 1458 GAAGTCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1517
Db 3327 TAAGTCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3386
Qy 1518 AGGATCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1577
Db 3387 AGGATCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3446
Qy 1578 CAGCAACAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1637
Db 3447 CACAATCAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3506
Qy 1638 TGGTCTCCATCCAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1697
Db 3507 TGGTCTCCATCCAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3566
Qy 1698 CATGAGTCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTT 1757
Db 3567 AACTAGATCTGATGTTGAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTT 3626
Qy 1758 CTGCAACAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1817
Db 3627 CTGCAACAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3686
Qy 1818 TTGCTCTCAAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1859
Db 3687 CCTCTCTCAAGCAAGGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3728

RESULT 12
ID AAQ35251
XX AAQ35251 standard; cDNA; 5406 BP.
AC AAQ35251:
XX
DT 25-JUN-1993 (first entry)
XX
DE Human flk-1 coding sequence.
XX

Human; receptor; protein; tyrosine kinase; pTK; primitive; mammalian; hematopoietic cell; pTK; mature; mHC; fetal; liver kinase 2; flk-1; liver; spleen; thymus; adult; brain; marrow; thymocyte; subset; multipotential; T-lymphoid; lineage; stomach; kidney; lung; heart; intestine; muscle; lymph node; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FI CDS 208..4311
FI /*tag= a
XX KW93000349-A.
XX 07-JAN-1993.
XX 26-JUN-1992: 92MO-US05401.
XX 28-JUN-1991: 91US-0728913.
XX 15-NOV-1991: 91US-0793065.
XX 24-DEC-1991: 91US-0813593.
XX 02-APR-1992: 92MO-US02750.
XX (UWPR-) UNIV PRINCETON.
XX Lemischka JB.
XX WPI: 1998-036323/04.
XX P-FSDS: AAR31377.
XX Nucleic acid encoding receptor protein tyrosine kinase - allows development of ligands to stimulate proliferation and/or differentiation of mammalian hematopoietic stem cells
XX Claim 10: Fig 2: 78pp: English.
XX This sequence encodes a human receptor protein tyrosine kinase which belongs to a new functional class of protein tyrosine kinases (pTKs). However, this pTK is not in the same class as flk-2. pTKs in the same class as flk-2 are expressed in primitive mammalian hematopoietic (pHC) cells but not in mature hematopoietic cells (mHC). This gene is expressed in more mature hematopoietic cells. The protein encoded by this sequence is an example of a receptor pTK and is called fetal liver kinase 1 (flk-1). flk-1 is expressed in fetal liver, spleen, thymus, brain, stomach, kidney, lung, heart and intestine, and adult brain, kidney, heart, spleen, lung, muscle, marrow and lymph nodes.
XX Sequence 5406 BP: 1412 A: 1299 C: 1422 G: 1273 T: 0 other;
Query Match 33.7%; Score 708.4; DB 14: Length 5406;
Seqs Local Similarity 65.0%; Pred. No. 1.3e-204;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;
Qy 201 CACTCTTAATCTACCAATCATGATTTCTCTGCAAGATTCAGGCACTATGCTTCAG 260
Db 2070 CATCTTGAATGTGSCATTTTCAAGATTCCTCTCTGAGGCAAGGCACTATGTTCTC 2129
Qy 261 AGCCAGGAATGTATACAGAGGGAAGAAATCTCCAGAGGAAGAAATTAACAATCAG 320
Db 2130 TCTTCAGATAGAGAGCAACAGAAAGACATTCCTGGTCAACAGCTATCATCTTAG 2189
Qy 321 TCAGGAAGCAACCACTCTCTCTGCAAACTCAGTATCACAGATGGCCATCAGCAGTTC 380
Db 2190 GGGATGSCACCCATGATCAGCGGAATCTGGAGATCAGCAACAACCAATTCGGAGAC 2249
Qy 381 CACCACTTTAGATGTCATGCTTCTCTGCAAACTCAGTATCACAGATGGCCATCAGCAGTTC 440
Db 2250 CATTTGAAGTGCATTCGCGGCACTCTGGAATCTTACCCACACATTCATGTTTCAAGA 2309
Qy 441 CAACCAAAAATACAGAGAGCTGGAATTTATTTAGGACCAAGGAGCAGCAGCTGTT 500
Db 2310 CAACGAGACCTTGGTAGAGATTCAGGCAATTTCTACTGAGAGATGGAGACCTGAC 2369

QY 501 TATTGAAGAGTACACAGAGAGATGAAGGTGTCTATACATGCAAGCCACACAGAA 560
 DB 2370 TATCCGAGAGTGAAGAGAGATGAAGGCTCTACACCTGCCAGGCTCAAGTCTC 2429
 QY 561 GGGTCTGTGGAGAGTGTAGCAATACCTCAGTGTCAAGAACCTCGGACAGGTCTAATCT 620
 DB 2430 TGGCTGTCAAGAGCGGACACCTCTTCAATAGAGGTGCCAGGAAAGACACAT 2489
 QY 621 GGAGTGTATCACTCAACAGCACCTGTGTGGGTGGGACTCTCTTCTGGCTCCTATTAAC 580
 DB 2490 GGAAGTCATTAATCTCTGGGCTGCACTGCACTGATTCGCTCTCTCTCTCTCT 2549
 QY 681 CTTCTTATCCGAATGAAGATC---TCTCTGAAATGAAGCTGACTACCTATC 737
 DB 2550 CATCTCTGAGGACCTTAAAGCGGCCAATGAAGGGGAACTGAAGACGCTACTTCTC 2609
 QY 738 AATTATATGACCCAGATGAAGTCTCTTGGATGAGCAGTGTGAGCGCTCTCTTATGA 757
 DB 2610 TATGTCAATGATCCAGATGANTGCCCTTGATGACCGCTGTGAACCTTGCCCTT 2659
 QY 798 TCCAGCAAGTGGAGTTTGCCTGGGAGAGACTTAACTGGGCAATCACTTGGAGAGG 857
 DB 2670 TCCAGCAAGTGGAGTTTCCAGGAGCGGCTGAAACTAGSAAACCTCTTGGCCGCG 2729
 QY 858 GCTTTTGAAGAGTGTTCACGATCAGCAATTTGGCAITAGAAATCACTTACGTGCGG 917
 DB 2730 TGCCTTGGCCAGATGATGAGCAGACGCTTTTGGATTTGACACAGACGACCTTGA 2789
 QY 918 GACTGTGGCTGTGAAATGCTGAAGAGGGGCCACGCGGAGTGAAGAAAGCTCTGAT 977
 DB 2790 AACAGTACCGCTCAAGATGTGAAGAGGAGCAACACAGACGAGCATCGAGCCCTCAT 2849
 QY 978 GACTGAGTAAATATCTTGACCCACATGCGCCACCATCTGAAGTGGTATACCTCTGGG 1037
 DB 2850 GTCTGAATCAAGATCTCAATCCACATGTTGTCACCATCTCAATGTGGTGAACCTCTAGG 2909
 QY 1038 AGCTTGACCAAGCAAGAGCGGCTCTGATGCTGATTTGAAATAGTCAATATGGA 1097
 DB 2910 CGCTTGACCAAGCGGAGGCGCTCTCATGGTATGTGGATCTCGAATTTGGAA 2969
 QY 1098 TCTTCCAACTACCTCAAGAGAAAGCTGACTTATTTTCTCAACAGAGTGCACACT 1157
 DB 2970 CTTATCAACTTACTTACGGGGCAAGAGAAATGAATTTGTTCCCTATAAGAGCAAGGGCG 3029
 QY 1158 ACAGATGGGCTTAAGAGAAATGAAGCCAGCGCTGGACAGGCAAGCAAGACCAAG 1217
 DB 3030 AGCTTCCGCGAGGCAAGCACTACCTTGG---GAGCTCTCGGTGATCTGAAAGAGC 3086
 QY 1218 ACTAGATAGCTACCAAGCAAGCAAGCTTTGCGAGCTCCGCTTTTCAAGAGCAATAAAG 1277
 DB 3087 CTGGAGAGCATACCAAGCAAGCAAGCTCTGCCAGCTCAGGCTTTTGTGGAGCAATC 3146
 QY 1278 TCTGAGTATGTTGAGAGAGAGGATCTGAGGTTTCTACAGAGGCCATCACTAT 1337
 DB 3147 GCTCAGTATGAGAGAGAGGATCTGAGAGGATCTGAGAGGATCTGAGAGGATCT 3206
 QY 1338 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGATGAGTCTGCTCTCCAG 1397
 DB 3207 GGAGATCTCATCTGTACAGTTTCAAGTGGCTTACAGGAGGATCTTCTGCACTCAG 3266
 QY 1398 AAGTGCATCTATCGGACCTGGCAGCAGAGAACTCTTTTATCTGAGAAACAGTGGT 1457
 DB 3267 GAAGTGTATCCAGAGGAGGACCTGGCAGCAGCAACATCTCTATCGGAGAGAGTGGT 3326
 QY 1458 GAGATTTGTGATTTGGCTTGGCGGATTTTATAGAACCCGATTTATGTGAGAA 1517
 DB 3327 TAAGATCTGTGATCTGCGCTTGGCCGAGGACATTTATAAGACCCGATTTATGTGAGAA 3386
 QY 1518 AGGAGATCTGCACTCTCTGAAATGGATGGCTCCCGAATCTATCTTTGACAAATCTA 1577
 DB 3387 AGAGATGCGGATCTCTCTTGAAGTGGATGGCCCGGAGAACTTTTTCAGAGATATA 3446
 QY 1578 CAGCACCAGAGCGAGCTGTGGTCTTACGGAGTATGCTGTGGGAAATCTCTCTCTAGG 1637

DB 3447 CACAATTCAGAGGATGTGGTCTTTCGGTGTGTGCTCTGGAAATATTTTCCTTAGG 3506
 QY 1638 TGGGTCTCCATACCCAGGAGTACAAATGGATGAGGACTTTTGCAGTCTGCTGAGGAGG 1697
 DB 3507 TGGCTTCCCATACCTGGGTCAAGATTTGATGAAGATTTTGTAGAGATTTGAAGAAGG 3566
 QY 1698 CATGAGATGAGAGCTCTCTGAGTACTTACCTTCCGAAATCTATCAGATCATCTGGACTG 1757
 DB 3557 AACTAGAAATGGGCTCTCTGACTACATACCCAGAAATGTACAGACCATTTGGGACTG 3626
 QY 1758 CTGGCAGAGAGCCCAAAAGAAAGCCCAAGATTTGCAGAACTTGTGAAAACTAGGTGA 1817
 DB 3627 CTGGCAGAGAGCCCAAAAGAAAGCCCAAGATTTGCAGAACTTGTGAAAGACTATAT 3686
 QY 1818 TTGCTTCAAGCAATGTACAAAGAGATGGTAAAGACTACAT 1859
 DB 3687 CCTCTGCAAGAAATGGCAGCAGGATGGCAAGACTATAT 3728
 RESULT 13
 ID A240916 standard: cDNA: 5406 BP.
 XX A240916
 AC A240916
 XX 12-OCT-1993 (first entry)
 DI Murine flk-1 cDNA.
 XX
 DE Murine; receptor; protein; tyrosine kinase; ptk; flk-1; flk-2; adult;
 KW family; primitive; hematopoietic cell; mature; fetal; liver; spleen;
 KW thymus; brain; stomach; kidney; lung; heart; intestine; bone marrow;
 KW muscle; lymph node; ss
 XX Mus musculus.
 XX
 XX Key Location/Qualifiers
 FH 208..4311
 FT CDS /*tag- a
 FT 208..264
 FT 265..4308
 FT mat_peptide /note- "Hydrophobic leader sequence"
 FT /*tag- c
 XX
 XX NC9310136-A.
 XX
 XX 27-MAY-1993.
 XX
 XX 16-NOV-1992: 92WO-US09893.
 XX
 XX 15-NOV-1991: 91US-0793065.
 XX (TYPR-) UNIV PRINCETON.
 XX
 XX Lemischka IR:
 XX
 XX NP1: 1993-182479/22.
 XX P-FSDB: AAB37504.
 XX
 XX Tcipoent hematopoietic stem cell receptors, their ligands and
 XX DNA sequences - for treating anaemia(s) and bone marrow damage
 XX due to e.g. cancer chemotherapy or radiotherapy
 XX
 XX Claim 14: Fig 2: 127pp; English.
 XX
 XX This sequence encodes the murine receptor protein tyrosine kinase
 XX (PTK), flk-1. This PTK is not in the same family as flk-2 (see also
 XX A240914-15) as it is expressed in primitive hematopoietic cells and
 XX also in mature hematopoietic cells. flk-1 is expressed in fetal
 XX liver, spleen, thymus, brain, stomach, kidney, lung, heart and
 XX intestine and adult brain, bone marrow, kidney, heart, spleen, lung,
 XX

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 09:21:42 : Search time 2723.29 seconds
(without alignments)
12494.717 Million cell updates/sec

Title: X51602_COPY_1900_4000

Perfect score: 2101

Sequence: 1 TATATACAGATGTGCCAAA.....CGACAGCAGCAGCIGIGG 2101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:**

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_est1:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pin:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_other:**

26: em_gss_pro:**

27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	827.6	39.4	3976	11 AK004721	Mus muscu
2	715	34.0	936	13 B1762725	B1762725 603048490
3	571.4	27.2	697	14 BQ604389	BQ604389 MI-P-Cp1-
4	561.4	26.7	666	9 AL042601	AL042601 DKF2p434L
5	558.2	26.6	650	12 BG005458	BG005458 MR3-GN018
6	466.8	22.2	1000	9 AL575294	AL575294 AL575294

7	451.6	21.5	566	14 BQ770907	BQ770907 UI-M-F10-
8	444.6	21.2	1005	9 AL570084	AL570084 AL570084
9	442.4	21.1	940	9 AL569839	AL569839 AL569839
10	423.4	20.2	855	9 A1326556	A1326556 mm09f03.y
11	415.2	19.8	512	12 BG384715	BG384715 303951.MA
12	406	19.3	606	13 B1458691	B1458691 603199829
13	385.6	18.4	561	9 AA066758	AA066758 mm09f03.r
14	377.4	18.0	527	9 AL571839	AL571839 AL571839
15	374.8	17.8	835	13 B1731060	B1731060 603351683
16	370.6	17.6	441	12 BG384720	BG384720 303961.MA
17	349.2	16.6	387	12 BF086406	BF086406 RC6-GN007
18	336.6	16.0	383	14 BQ387353	BQ387353 PM4-GN030
19	332.4	15.8	624	12 BG435852	BG435852 602508476
20	321.6	15.3	500	14 R62703	R62703 Y110b03.r1
21	321.2	15.3	645	10 AW167792	AW167792 XG56D06.x
22	321.2	15.3	652	9 A1800102	A1800102 TT20F07.x
23	321.2	15.3	782	9 AL573250	AL573250 AL573250
24	321.2	15.3	797	12 BG571885	BG571885 602593233
25	320.8	15.3	955	9 AL576447	AL576447 AL576447
26	315.8	15.0	689	13 BQ075253	BQ075253 B3075253
27	312.2	14.9	656	10 BB577273	BB577273 BB577273
28	308	14.7	859	9 AL573279	AL573279 AL573279
29	306	14.6	581	9 A194109	A194109 Z138407.s
30	303	14.4	496	10 AW129059	AW129059 XE94f10.x
31	295.8	14.1	561	10 AW026025	AW026025 WY98f10.x
32	291	13.9	623	12 BF087145	BF087145 PM1-HT045
33	288.2	13.7	711	9 AA880212	AA880212 VX39D08.r
34	280.6	13.4	819	9 AL570470	AL570470 AL570470
35	278.8	13.3	672	9 A1761753	A1761753 W67H06.x
36	277.6	13.2	304	12 BG015691	BG015691 CM1-GN036
37	277.6	13.2	1067	14 BM925760	BM925760 AGENCOURT
38	274.2	13.1	681	14 W26612	W26612 36d4 Human
39	267	12.7	463	10 AW681286	AW681286 EST00004
40	266	12.7	671	9 A1568689	A1568689 AL568689
41	265.2	12.6	504	9 A1089100	A1089100 C082q10.s
42	264.6	12.6	472	12 BF061039	BF061039 7162h11.x
43	260.4	12.4	547	10 BE030679	BE030679 128743.MA
44	259.6	12.4	320	9 A1716140	A1716140 UI-R-Y0-a
45	257.6	12.3	359	12 BF543282	BF543282 UI-R-Y0-a

ALIGNMENTS

RESULT 1	AK004721	3976 bp	mkna	linear	HIC 19-JAN-2002
LOCUS	AK004721	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200012B22:FMS-like tyrosine kinase 1, full insert sequence.			
DEFINITION	AK004721.2	GI:15919895			
ACCESSION	AK004721				
VERSION	AK004721.2	GI:15919895			
KEYWORDS	HIC: CAP trapper.				
SOURCE	Mus musculus (strain:CS7BL/6J) adult male lung cDNA to mRNA, clone:1200012B22.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	1 Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10345636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Suganara, Y., Shibata, K., Itoh, M., Kondo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagacka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunagi, I., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, I., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, I., Matsuda, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multi-capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I.,
Saito, T., Okazaki, Y., Gojobori, I., Bono, H., Kasukawa, I., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, R.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okada, T., Furuno, M., Kono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Noradre, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Tato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J. H., Kotsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 (bases 1 to 3976)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hayashizaki, Y.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F.,
Hume, D., Imolani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, I.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyama, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ohno, M., Okazaki, Y., Okada, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasai, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, I.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, I., Yamakawa, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.
Direct submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN) Laboratory for Genomes
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Sucho-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-5222,
Fax: 81-45-503-9216)
On Oct 4, 2001 this sequence version replaced gi:12836104.
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGACGAGGCGCGCCGACGCTGTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence [5'

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

6 (bases 1 to 3976)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hayashizaki, Y.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F.,
Hume, D., Imolani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, I.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyama, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ohno, M., Okazaki, Y., Okada, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasai, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, I.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, I., Yamakawa, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.
Direct submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN) Laboratory for Genomes
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Sucho-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-5222,
Fax: 81-45-503-9216)
On Oct 4, 2001 this sequence version replaced gi:12836104.
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGACGAGGCGCGCCGACGCTGTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence [5'

COMMENT

12836104
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGACGAGGCGCGCCGACGCTGTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence [5'

GAGACGAGGCGCGCCGACGCTGTTTCTTTTCTTNN 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end, SstI; 3' end:
XhoI. Host: SOLR.
Location/Qualifiers
Source
1. 3976
/organism="Mus musculus"
/strain="C57BL/6J"
/3d_xref="FANTOM_DB:120001B22"
/db_xref="MGD:MGI:1904343"
/db_xref="taxon:10090"
/clone="120001B22"
/sex="male"
/issue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. 3976
/gene="Ftl1"
/gene="Ftl1"
/gene="Ftl1"
/note="FMS-like tyrosine kinase 1
data source:MGD, source key:MGI:95558, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB23504.2"
/db_xref="GI:15915896"
/db_xref="MGD:MGI:95558"
/translation="KILIHGHLNVLNLSGKGGPLMIVVEYKYNLSNLS
KADFLNKAAALHMEIKLEPELEQKPRLLSVSSSVTSFPPEDRSVDVPG
DEDSEYKQPLIMEDLISYSPQVARGMEFLSKKCIHRDLAARNILLSENWVYDVG
FGLARDYKPNPVRKGRDRLPLKNAPESIFDKVSTKSDVMSGVLLWIFSLGGS
PNPQVQDQCKSLKQGMRTPEVATPEYVQIMLECKDKPKERKPEFAFLVZKLDG
LIQANVQDQCKDPIPLNAILTNSGTFYTFSELLFKDGFADFFHFGSSDDRVY
NAPFMLEKIKFIEELSPNSFMFQGLDITSLISPLLKRFETWIKFKASMKID
LNFASKSEAGLSLDLPSPFCFSSOCHIRPVDSELSKSCSPPPDPVNSVVLVSS
PPA."

FEATURES
Source

1. 3976
/organism="Mus musculus"
/strain="C57BL/6J"
/3d_xref="FANTOM_DB:120001B22"
/db_xref="MGD:MGI:1904343"
/db_xref="taxon:10090"
/clone="120001B22"
/sex="male"
/issue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. 3976
/gene="Ftl1"
/gene="Ftl1"
/gene="Ftl1"
/note="FMS-like tyrosine kinase 1
data source:MGD, source key:MGI:95558, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB23504.2"
/db_xref="GI:15915896"
/db_xref="MGD:MGI:95558"
/translation="KILIHGHLNVLNLSGKGGPLMIVVEYKYNLSNLS
KADFLNKAAALHMEIKLEPELEQKPRLLSVSSSVTSFPPEDRSVDVPG
DEDSEYKQPLIMEDLISYSPQVARGMEFLSKKCIHRDLAARNILLSENWVYDVG
FGLARDYKPNPVRKGRDRLPLKNAPESIFDKVSTKSDVMSGVLLWIFSLGGS
PNPQVQDQCKSLKQGMRTPEVATPEYVQIMLECKDKPKERKPEFAFLVZKLDG
LIQANVQDQCKDPIPLNAILTNSGTFYTFSELLFKDGFADFFHFGSSDDRVY
NAPFMLEKIKFIEELSPNSFMFQGLDITSLISPLLKRFETWIKFKASMKID
LNFASKSEAGLSLDLPSPFCFSSOCHIRPVDSELSKSCSPPPDPVNSVVLVSS
PPA."

gene

1. 3976
/gene="Ftl1"
/gene="Ftl1"
/gene="Ftl1"
/note="FMS-like tyrosine kinase 1
data source:MGD, source key:MGI:95558, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB23504.2"
/db_xref="GI:15915896"
/db_xref="MGD:MGI:95558"
/translation="KILIHGHLNVLNLSGKGGPLMIVVEYKYNLSNLS
KADFLNKAAALHMEIKLEPELEQKPRLLSVSSSVTSFPPEDRSVDVPG
DEDSEYKQPLIMEDLISYSPQVARGMEFLSKKCIHRDLAARNILLSENWVYDVG
FGLARDYKPNPVRKGRDRLPLKNAPESIFDKVSTKSDVMSGVLLWIFSLGGS
PNPQVQDQCKSLKQGMRTPEVATPEYVQIMLECKDKPKERKPEFAFLVZKLDG
LIQANVQDQCKDPIPLNAILTNSGTFYTFSELLFKDGFADFFHFGSSDDRVY
NAPFMLEKIKFIEELSPNSFMFQGLDITSLISPLLKRFETWIKFKASMKID
LNFASKSEAGLSLDLPSPFCFSSOCHIRPVDSELSKSCSPPPDPVNSVVLVSS
PPA."

CDs

1. 3976
/gene="Ftl1"
/gene="Ftl1"
/gene="Ftl1"
/note="FMS-like tyrosine kinase 1
data source:MGD, source key:MGI:95558, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB23504.2"
/db_xref="GI:15915896"
/db_xref="MGD:MGI:95558"
/translation="KILIHGHLNVLNLSGKGGPLMIVVEYKYNLSNLS
KADFLNKAAALHMEIKLEPELEQKPRLLSVSSSVTSFPPEDRSVDVPG
DEDSEYKQPLIMEDLISYSPQVARGMEFLSKKCIHRDLAARNILLSENWVYDVG
FGLARDYKPNPVRKGRDRLPLKNAPESIFDKVSTKSDVMSGVLLWIFSLGGS
PNPQVQDQCKSLKQGMRTPEVATPEYVQIMLECKDKPKERKPEFAFLVZKLDG
LIQANVQDQCKDPIPLNAILTNSGTFYTFSELLFKDGFADFFHFGSSDDRVY
NAPFMLEKIKFIEELSPNSFMFQGLDITSLISPLLKRFETWIKFKASMKID
LNFASKSEAGLSLDLPSPFCFSSOCHIRPVDSELSKSCSPPPDPVNSVVLVSS
PPA."

BASE COUNT

1058 a 950 c 598 g 970 t
ORIGIN
Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

ORIGIN

Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

ORIGIN

Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

ORIGIN

Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

ORIGIN

Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

ORIGIN

Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

ORIGIN

Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

ORIGIN

Query match 39.48; Score 827.6; DB li: Length 3976;
Best local similarity 84.68; Pred. No. 1.7e-223;
Matches 942; Conservative 0; Mismatches 165; Gaps 1;
Query 368 AAAATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 1047
DB 1 AAGATCTTGACCCACATTCGACCATCTGACGGTGGTTACCTGCTGGAGGCTGCAC 60

[illegible]

RESULT 2					
B1762725	LOCUS	B1762725	936 bp	mRNA	linear EST 25-SEP-2001
	DEFINITION	603048490P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:518859 5'			
					mRNA sequence.

PD 15-JUN-1999.
 XX 10-FEB-1998; 98US-0021324.
 XX 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252498.
 PR 15-FEB-1996; 96US-0601891.

(UYPR-) UNIV PRINCETON.

Lemischka IR;

WPI; 1999-357194/30.
 P-PSDB; AAY08618.

Isolating hematopoietic cells expressing fetal liver kinase 1
 receptors

Disclosure; Fig 2; 59pp; English.

This invention describes a novel method of isolating cells expressing fetal liver kinase 1 (flk-1) receptors on their surface and comprises binding the cells to a polyclonal or monoclonal antibody specific to the flk-1 receptor and isolating the cells that have bound to the antibody. The method can be used to isolate hematopoietic stem cells in any mammal but preferably a rat, mouse, rabbit or human. The proteins of the invention belong to the receptor protein family. This sequence encodes the murine flk-1 protein which is used in the method of the invention.

SQ Sequence 5404 BP; 1411 A; 1297 C; 1423 G; 1273 T; 0 other;

Query Match 24.8%; Score 74.6; DB 20; Length 5404;
 Best Local Similarity 78.8%; Pred. No. 1.8e-13;
 Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTTCTCTGAATGGATGGCTCTGATCTATCTTTGACAAAATCTACAGCACCA 128
 DB 3395 CCCGACTCCCTTTGAAGTGGATGGCCCGGAAACCATTTTTCACAGATATACACATTC 3454
 129 AGAGCGACGTGTGCTCTTACGAGTATCTGCTGGGAATCTTCTCCTTAGGT 181
 3455 AGAGCGATGTGCTCTTTCGGTGTGCTGCTGGGAATATTTTCTCCTTAGGT 3507

RESULT 12

AAQ29957
 ID AAQ29957 standard; cDNA; 5406 BP.

AC AAQ29957;

DT 15-MAR-1993 (first entry)

DE flk-1 cDNA sequence.

Thymidine kinase; TK; hematopoietic; stem cells; proliferation;
 differentiation; progenitor cells; foetal liver kinase; ss.

OS Mus musculus.

Key Location/Qualifiers
 CDS 208..4311
 /tag= a

PN W09217486-A.

PD 15-OCT-1992.
 XX 02-APR-1992; 92WC-US02750.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 XX (UYPR-) UNIV PRINCETON.
 XX Lemischka IR;
 DR WPI; 1992-366185/44.
 DR P-PSDB; AAR28041.
 XX Stimulating proliferation and/or differentiation of primitive
 mammalian hematopoietic stem cells - using ligand that binds
 thymidine kinase and flk-1 and flk-2
 XX Claim 14; Fig 2; 94pp; English.

The murine foetal liver kinase (flk-1) clone was isolated by standard PCR techniques from stem-cell receptor-contg. tissue cDNA libraries. Suitable tissues include foetal liver, spleen or thymus cells or adult marrow or brain cells. The PCR primers used are based on known sections of the flk-1 gene. The murine flk-1 clone may be used in a vector to transform hematopoietic cells. The thymidine kinase encoded by flk-1 is expressed in primitive but not mature hematopoietic cells. Ligand binding to the TK may be prepd. which can stimulate proliferation and/or differentiation of primitive hematopoietic cells in vivo. The ligands can stimulate the proliferation of additional primitive stem cells, differentiation into more mature progenitor cells, or both.

See also AA029954-6.

SQ Sequence 5406 BP; 1412 A; 1294 C; 1427 G; 1273 T; 0 other;

Query Match 24.8%; Score 74.6; DB 13; Length 5406;
 Best Local Similarity 78.8%; Pred. No. 1.8e-13;
 Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTTCTCTGAATGGATGGCTCTGATCTATCTTTGACAAAATCTACAGCACCA 128
 DB 3395 CCCGACTCCCTTTGAAGTGGATGGCCCGGAAACCATTTTTCACAGATATACACATTC 3454
 129 AGAGCGACGTGTGCTCTTACGAGTATCTGCTGGGAATCTTCTCCTTAGGT 181
 3455 AGAGCGATGTGCTCTTTCGGTGTGCTGCTGGGAATATTTTCTCCTTAGGT 3507

RESULT 13

AAQ53504
 ID AAQ53504 standard; cDNA; 5406 BP.

AC AAQ53504;

DT 27-JUN-1994 (first entry)

DE Murine flk-1 cDNA.

Receptor protein tyrosine kinase; pTK family; foetal liver kinase;
 mfk; primitive; totipotent; hematopoietic cell; stem cell;
 proliferation; stromal cell; ds.

OS Mus musculus.

Key Location/Qualifiers
 CDS 208..4311
 /tag= a

FT sig_peptide

FT /product= mfk-1
 FT 208..264
 FT /tag= b

FT mat_peptide /note= "hydrophobic leader"
 FT 265..4308
 FT /*tag= c

US5270458-A.

14-DEC-1993.

02-APR-1991; 91US-0679666.

02-APR-1991; 91US-0679666.

28-JUN-1991; 91US-0728913.

15-NOV-1991; 91US-0793065.

24-DEC-1991; 91US-0813593.

26-JUN-1992; 92US-0906397.

12-NOV-1992; 92US-0975049.

19-NOV-1992; 92US-0977451.

(UYPR-) UNIV PRINCETON.

Lemischka IR;

WPI; 1993-405021/50.

P-PSDB; AAR44996.

Isolated nucleic acid molecules of hematopoietic stem cell

receptor flk-2 - encoding mammalian receptor protein tyrosine

kinases expressed in primitive haematopoietic cells

Disclosure; Fig 2; 60pp; English.

Nucleic acid sequences coding for murine flk-2 and specified

subfragments of it are claimed. The murine flk-1 coding sequence

(i.e. A053504) is also disclosed. The flk polypeptides are

receptor protein tyrosine kinases which are expressed only in

primitive haematopoietic cells.

Sequence 5406 BP; 1412 A; 1298 C; 1423 G; 1273 T; 0 other;

Query Match 24.8%; Score 74.6; DB 14; Length 5406;

Best Local Similarity 78.8%; Pred. No. 1.8e-13;

Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTTCCTTGAAATGGATGGCTCTGAATCTATCTTGACAAATCTACAGACCA 128

Db 3395 CCCGACTCCCTTTGAAGTGGATGGCCCCGGAACCAATTTTGACAGATATACAAATTC 3454

QY 129 AGACGACGTGGTCTTACGAGATATGCTGTGGAAATCTTCCTTAGGT 181

3455 AGACGATGTGGTCTTTCGGTGTGTTGCTGTGGAAATATTTTCCTTAGGT 3507

RESULT 14

AAQ35251

ID AAQ35251 standard; cDNA; 5406 BP.

XX AC AAQ35251;

XX 25-JUN-1993 (first entry)

XX Human flk-1 coding sequence.

XX Human; receptor; protein; tyrosine kinase; pTK; primitive; mammalian;

XX hematopoietic cell; pHC; mature; mHC; fetal; liver kinase 2; flk-1;

XX liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;

XX multipotential; T-lymphoid; lineage; stomach; kidney; lung; heart;

XX intestines; muscle; lymph node; ss.

XX WC9300349-A.

XX 07-JAN-1993.

XX 26-JUN-1992; 92WC-US05401.

XX 26-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 02-APR-1992; 92WC-US02750.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1993-036323/04.

XX P-PSDB; AAR31377.

XX Nucleic acid encoding receptor protein tyrosine kinase - allows

development of ligands to stimulate proliferation and/or

differentiation of mammalian haematopoietic stem cells

Claim 10; Fig 2; 78pp; English.

XX This sequence encodes a human receptor protein tyrosine kinase which

belongs to a new functional class of protein tyrosine kinases (pTKs).

However, this pTK is not in the same class as flk-2. pTKs in the same

class as flk-2 are expressed in primitive mammalian hematopoietic

(pHC) cells but not in mature hematopoietic cells (mHC). This gene

is expressed in more mature hematopoietic cells. The protein encoded

by this sequence is an example of a receptor pTK and is called fetal

liver kinase 1 (flk-1). flk-1 is expressed in fetal liver/spleen,

thymus, brain, stomach, kidney, lung, heart and intestine, and adult

brain, kidney, heart, spleen, lung, muscle, marrow and lymph nodes.

XX Sequence 5406 BP; 1412 A; 1299 C; 1422 G; 1273 T; 0 other;

Query Match 24.8%; Score 74.6; DB 14; Length 5406;

Best Local Similarity 78.8%; Pred. No. 1.8e-13;

Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTTCCTTGAAATGGATGGCTCTGAATCTATCTTGACAAATCTACAGACCA 128

Db 3395 CCCGACTCCCTTTGAAGTGGATGGCCCCGGAACCAATTTTGACAGATATACAAATTC 3454

QY 129 AGACGACGTGGTCTTACGAGATATGCTGTGGAAATCTTCCTTAGGT 181

3455 AGACGATGTGGTCTTTCGGTGTGTTGCTGTGGAAATATTTTCCTTAGGT 3507

RESULT 15

AAQ40516

ID AAQ40516 standard; cDNA; 5406 BP.

XX AC AAQ40516;

XX 19-OCT-1993 (first entry)

XX Marine flk-1 cDNA.

XX Marine; receptor; protein; tyrosine kinase; pTK; flk-1; flk-2; adult;

XX family; primitive; hematopoietic cell; mature; fetal; liver; spleen;

XX thymus; brain; stomach; kidney; lung; heart; intestine; bone marrow;

XX muscle; lymph node; ss.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 208..4311

XX /*tag= a

XX sig_peptide 208..264

XX /*tag= b

FT mat_peptide /note= "Hydrophobic leader sequence"
 FT 265..4308
 FT /*tag= C

XX WO9310136-A.

XX 27-MAY-1993.

XX 16-NOV-1992; 92WO-US09893.

XX 15-NOV-1991; 91US-0793065.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI: 1993-182479/22.

XX P-PSDB; AAR37504.

XX Totipotent haematopoietic stem cell receptors, their ligands and
 XX DNA sequences - for treating anaemia(s) and bone marrow damage
 XX due to e.g. cancer chemotherapy or radiotherapy

XX Claim 14; Fig 2; 127pp; English.

XX This sequence encodes the murine receptor protein tyrosine kinase
 XX (pTK), flk-1. This pTK is not in the same family as flk-2 (see also
 XX AAQ0914-15) as it is expressed in primitive hematopoietic cells and
 XX also in mature hematopoietic cells. flk-1 is expressed in fetal
 XX liver, spleen, thymus, brain, stomach, kidney, lung, heart and
 XX intestine and adult brain, bone marrow, kidney, heart, spleen, lung,
 XX muscle and lymph nodes.

SQ Sequence 5406 BP; 1412 A; 1298 C; 1423 G; 1273 T; 0 other;

Query Match 24.8%; Score 74.6; DB 14; Length 5406;
 Best Local Similarity 78.8%; Pred. No. 1.8e-13;
 Matches 59; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTCCCTGAAATGGATGCTCTGATCTATCTTTGACAAATCTACAGCACCA 128
 | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

Db 3395 CCGGACTCCCTTGAAGTGGATGGCCCGGAAACCAATTTTGACACAGTATACAAATTC 3454

QY 129 AGAGCGACCTGCTCTTACGAGATGCTGTGGGAAATCTTCCTTAGGT 181
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 3455 AGAGCGATGTGCTCTTCGGTGTGTGCTCTGGGAAATATTTCCTTAGGT 3507

Search completed: December 8, 2002, 06:01:53
 Job time : 360.95 secs

GenCore version 5.1.3
Copyright (C) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 05:16:46 : Search time 2350.03 Seconds
(without alignments)
3665.203 Million cell updates/sec

Title: US-09-778-900a-3_COPY_200_500

Perfect score: 301

Sequence: 1 ggacactcgggtgaatg.....taacaaaaactcttaagt 301

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

tal number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_ov:*

7: gb_ph:*

8: gb_pi:*

9: gb_pi:*

10: gb_pi:*

11: gb_pi:*

12: gb_pi:*

13: gb_pi:*

14: gb_pi:*

15: gb_pi:*

16: gb_pi:*

17: gb_pi:*

18: gb_pi:*

19: gb_pi:*

20: gb_pi:*

21: gb_pi:*

22: gb_pi:*

23: gb_pi:*

24: gb_pi:*

25: gb_pi:*

26: gb_pi:*

27: gb_pi:*

28: gb_pi:*

29: gb_pi:*

30: gb_pi:*

31: gb_pi:*

32: gb_pi:*

33: gb_pi:*

34: gb_pi:*

35: gb_pi:*

36: gb_pi:*

37: gb_pi:*

38: gb_pi:*

39: gb_pi:*

40: gb_pi:*

41: gb_pi:*

42: gb_pi:*

43: gb_pi:*

44: gb_pi:*

45: gb_pi:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	299.4	99.5	164519	9	AL138712 Human DNA
C 2	124.2	41.3	101987	2	AC128808 Rattus no
C 3	124.2	41.3	144090	2	AC128409 Rattus no
C 4	114.4	38.0	4017	6	AX481481 Sequence
C 5	114.4	38.0	4017	9	AF063657 Homo sapi
C 6	112.8	37.5	7680	9	HSFLT
C 7	104.4	34.7	220965	2	HSAL12688
C 8	102.8	34.2	197279	9	AC104439
C 9	95.8	31.8	4734	10	RATFLT1
C 10	95.8	31.8	5482	10	MMEMRK2
C 11	95.8	31.8	6055	10	MUSRTKB
C 12	95.8	31.8	6275	10	D88689
C 13	94.2	31.3	666	4	AF233077
C 14	88	29.2	4272	5	AB065372
C 15	80.6	26.8	211330	10	AL646088
C 16	77.2	25.6	459	4	BTVEGFREC
C 17	76.2	25.3	4044	6	AX481480
C 18	76.2	25.3	4071	6	AR142803
C 19	76.2	25.3	4071	6	AR201382
C 20	76.2	25.3	4071	6	AX481483
C 21	76.2	25.3	4225	6	AX133539
C 22	76.2	25.3	4225	6	AX133539
C 23	76.2	25.3	4230	9	HSGFRK
C 24	76.2	25.3	4230	9	HUMKDRZ
C 25	76.2	25.3	4236	6	AF030758
C 26	76.2	25.3	5830	9	AF035121
C 27	75.4	25.0	163017	2	AC131433
C 28	74.6	24.8	4790	10	S53103
C 29	74.6	24.8	5391	10	MMFLK1
C 30	74.6	24.8	5406	6	AF005213
C 31	74.6	24.8	5406	6	AR071705
C 32	74.6	24.8	5406	6	I25171
C 33	74.6	24.8	5406	6	I40602
C 34	74.6	24.8	5464	10	BC020530
C 35	74.6	24.8	5470	6	AR068047
C 36	74.6	24.8	5470	6	AR125598
C 37	74.6	24.8	5470	10	MMFLK1M
C 38	74.2	24.7	73760	2	AC101367
C 39	74.2	24.7	76721	9	AL591024
C 40	74.2	24.7	209402	2	AC120348
C 41	74.2	24.7	264425	10	AC019028
C 42	73.6	24.5	5284	10	MUSRTKA
C 43	73	24.3	5892	10	ENUS306
C 44	72	23.9	1040	5	AF007760
C 45	72	23.9	4360	10	AF402785

ALIGNMENTS

RESULT 1	AL138712/c	AL138712	164519 bp	DNA	linear	PRI 03-AUG-2001
LOCUS	Human DNA sequence from clone RP11-502P18 on chromosome 13,	complete sequence.				
DEFINITION	Human DNA sequence from clone RP11-502P18 on chromosome 13,					
ACCESSION	AL138712	GI:15131448				
VERSION	AL138712					
KEYWORDS	HIG.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;					
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1 (bases 1 to 164519)					
TITLE	Bates, K.					
	Direct Submission					

umber of results predicted by chance to have a

JOURNAL

Submitted (03-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk. Clone requests: clonerequests@sanger.ac.uk

COMMENT

On Aug 9, 2001 this sequence version replaced g1:14715292. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone of more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; S, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; information on the WormPEP database can be found at

<http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13> Rpl1-502p18 is from the library RPCR-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone Rpl1-502p18. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true right end of clone Rpl1-502p18 is at 164519 in this sequence. The true left end of clone Rpl1-95G6 is at 140286 in this sequence. The true right end of clone Rpl1-179f17 is at 2000 in this sequence.

FEATURES

source

Location/Qualifiers

1. 164519
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="Rpl1-502p18"
/clone_lib="RPC1-11.2"
216..318
/note="D6 repeat: matches 2. 104 of consensus"
983..1279
/note="AluSg repeat: matches 1. 297 of consensus"
1634..1706
/note="MER5A repeat: matches 95. 168 of consensus"
1989..2053
/note="AluJ/FLAM repeat: matches 21. 86 of consensus"
2724..2892
/note="L2 repeat: matches 2206. 2398 of consensus"
3392..3435
/note="22 copies 2 mer ac 93% conserved"
3453..3743
/note="AluJ repeat: matches 1. 291 of consensus"
3982..4146
/note="MER5B repeat: matches 1. 161 of consensus"
4495..4586
/note="FLAM.C repeat: matches 33. 124 of consensus"
6847..6934
/note="44 copies 2 mer ca 79% conserved"
8591..9005
/note="AluJ repeat: matches 1. 312 of consensus"
10412..10773
/note="L1 repeat: matches 4244. 4621 of consensus"
10814..10984
/note="L1 repeat: matches 3996. 4169 of consensus"
11393..11698
/note="AluSg repeat: matches 1. 307 of consensus"
12456..12549

repeat_region
/note="L1M4 repeat: matches 4612. 4706 of consensus"
12550..12859
/note="AluSg repeat: matches 1. 310 of consensus"
12860..13131
/note="L1M4 repeat: matches 4344. 4612 of consensus"
13227..13760
/note="L1M7 repeat: matches 5629. 6169 of consensus"
13789..14015
/note="L1M7 repeat: matches 5250. 5477 of consensus"
14017..16233
/note="L1M4 repeat: matches 4063. 6299 of consensus"
16234..16506
/note="AluJ repeat: matches 30. 301 of consensus"
16507..16680
/note="L1M4 repeat: matches 3892. 4063 of consensus"
16689..16819
/note="L1M4 repeat: matches 5034. 5255 of consensus"
16910..17043
/note="L1M4 repeat: matches 4958. 5103 of consensus"
17160..17462
/note="AluSg repeat: matches 1. 302 of consensus"
17794..17914
/note="L1M7 repeat: matches 1622. 1739 of consensus"
17915..18088
/note="AluJ repeat: matches 144. 295 of consensus"
18089..18379
/note="AluSg repeat: matches 1. 292 of consensus"
18380..18516
/note="AluJ repeat: matches 1. 144 of consensus"
18517..18634
/note="L1M7 repeat: matches 1495. 1622 of consensus"
18637..18846
/note="105 copies 2 mer tt 57% conserved"
19590..19900
/note="AluSg repeat: matches 1. 311 of consensus"
19902..19923
/note="11 copies 2 mer tt 100% conserved"
21313..21486
/note="FAM repeat: matches 1. 175 of consensus"
21878..22123
/note="MIR repeat: matches 1. 251 of consensus"
24957..25516
/note="Sequence from clone PCR only."
25177..25343
/note="Charlie repeat: matches 48. 215 of consensus"
25374..25537
/note="82 copies 2 mer ta 61% conserved"
25538..25830
/note="AluSg repeat: matches 1. 268 of consensus"
26437..26754
/note="MER2 repeat: matches 2. 311 of consensus"
26845..27004
/note="FAM repeat: matches 1. 160 of consensus"
27014..27230
/note="AluSg/x repeat: matches 89. 305 of consensus"
27233..27286
/note="AluSg/x repeat: matches 79. 133 of consensus"
27339..27650
/note="AluSg repeat: matches 1. 313 of consensus"
27679..27810
/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
27828..28123
/note="AluSg repeat: matches 1. 299 of consensus"
29264..29564
/note="AluJ repeat: matches 1. 301 of consensus"
29773..29973
/note="MIR repeat: matches 37. 235 of consensus"
31059..31337
/note="AluJ repeat: matches 7. 289 of consensus"
31971..32253
/note="AluSg repeat: matches 2. 297 of consensus"

According to the Pre Publication Rules, every patent application received by the United States Patent and Trademark Office after November 29, 2000 will be pre-published at eighteen months from the effective filing date. When the application is published the contents, including the sequences, will become prior art.

Two new databases have been created to hold the pre-published sequences:

Published_Applications_NA contains nucleic acid sequences; the search results will have the extension **.rnpb**.

Published_Applications_AA contains amino acid sequences; the search results will have the extension **.rapb**.

Each pre-published application is given a unique Publication Number. An example of a Publication Number is US20021234567A1. The "US" indicates the application was a U.S. application. The first 4 digits show the calendar year the application was published. The next 7 digits represent when the application was published. This 7-digit number starts at zero at the beginning of each calendar year. Each application published is given the next number in order. The "A" indicates a utility patent application and the "1" shows that this was the first time the application had been published. If the applicants submit changes to the application, they may request that the changed application be published again. In such instances, the "1" at the end of the number would be replaced by a "2".

Sequences in the PGPub database are public information; it is permissible to leave these results in the case.

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 11:59:28 ; Search time 52.5871 Seconds
(without alignments)
5977.502 Million cell updates/sec

Title: D64016_COPY_500_1440

Perfect score: 941

Sequence: 1 GCGGAGAAAGACGACAC.....GCGGGGCTGTCGCGCGGG 941

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 350425 seqs, 19496369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA..*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq..*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq..*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq..*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq..*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq..*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq..*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq..*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq..*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq..*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq..*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq..*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq..*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq..*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq..*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.6	5.6	1064	10 US-09-804-682-29	Sequence 29, Appl
2	51.4	5.5	1919	10 US-09-814-777A-17	Sequence 17, Appl
3	50.6	5.4	43058	10 US-09-954-456-292	Sequence 292, Appl
4	50.6	5.4	43058	10 US-09-954-456-529	Sequence 529, Appl
5	50.6	5.4	43058	10 US-09-880-107-3950	Sequence 3950, Appl
6	48.4	5.1	42999	9 US-09-755-462A-17	Sequence 17, Appl
7	48.4	5.1	42999	9 US-10-125-767-17	Sequence 17, Appl
8	47	5.0	2561	12 US-10-023-523-48	Sequence 48, Appl
9	47	5.0	2561	12 US-10-023-523-48	Sequence 48, Appl
10	46.4	4.9	2481	10 US-09-894-968-35	Sequence 35, Appl
11	45	4.8	1614	12 US-10-023-523-45	Sequence 45, Appl
12	45	4.8	9968	10 US-09-764-877-2718	Sequence 2718, Appl
13	45	4.8	12425	12 US-10-023-523-50	Sequence 50, Appl
14	45	4.8	12425	12 US-10-023-523-50	Sequence 50, Appl
15	44.8	4.8	8459	10 US-09-817-913-8	Sequence 8, Appl
16	44.8	4.8	8459	10 US-09-817-913-8	Sequence 8, Appl
17	44.4	4.7	43058	10 US-09-954-456-292	Sequence 292, Appl
18	44.4	4.7	43058	10 US-09-954-456-292	Sequence 292, Appl
19	44.4	4.7	43058	10 US-09-954-456-529	Sequence 529, Appl

C 20	44.4	4.7	43058	10 US-09-880-107-3950	Sequence 3950, Appl
C 21	44.2	4.7	2236	9 US-09-978-295A-399	Sequence 399, Appl
C 22	44.2	4.7	2236	9 US-09-978-697-399	Sequence 399, Appl
C 23	44.2	4.7	2236	9 US-09-978-192A-399	Sequence 399, Appl
C 24	44.2	4.7	2435	10 US-09-823-038A-40	Sequence 40, Appl
C 25	43.4	4.6	8459	10 US-09-817-913-8	Sequence 8, Appl
C 26	43.4	4.6	8459	10 US-09-817-913-8	Sequence 8, Appl
C 27	43.4	4.6	152331	9 US-10-095-407-16	Sequence 16, Appl
C 28	43.2	4.6	6314	10 US-09-954-456-319	Sequence 319, Appl
C 29	43.2	4.6	6314	10 US-09-954-456-1230	Sequence 1230, Appl
C 30	43.2	4.5	15500	10 US-09-764-869-2109	Sequence 2109, Appl
C 31	42.6	4.5	3110	10 US-09-764-877-3912	Sequence 3912, Appl
C 32	42.6	4.5	3110	10 US-09-764-877-3914	Sequence 3914, Appl
C 33	42.6	4.5	4689	9 US-09-860-846-34	Sequence 34, Appl
C 34	42.6	4.5	4689	9 US-09-860-846-34	Sequence 34, Appl
C 35	42.6	4.5	36778	9 US-09-861-289-34	Sequence 5, Appl
C 36	42.6	4.5	36778	9 US-09-861-289-5	Sequence 5, Appl
C 37	42.6	4.5	174424	10 US-09-967-768A-314	Sequence 314, Appl
C 38	42.4	4.5	1065	10 US-09-804-682-33	Sequence 33, Appl
C 39	42.2	4.5	1887	9 US-09-954-456-1624	GENERAL INFORMATION
C 40	42.2	4.5	3547	9 US-10-044-716-7	Sequence 1, Appl
C 41	42.2	4.5	4348	10 US-09-879-445-1	Sequence 5, Appl
C 42	42.2	4.5	8147	12 US-10-109-886-9	Sequence 9, Appl
C 43	42.2	4.5	152331	9 US-10-095-407-16	Sequence 16, Appl
C 44	42	4.5	242	10 US-09-983-965-4700	Sequence 176, Appl
C 45	42	4.5	5025	10 US-09-960-253-176	

ALIGNMENTS

RESULT 1
US-09-804-682-29/c
Sequence 29, Application US/09804682
Patent No. US20020106763A1
GENERAL INFORMATION:
APPLICANT: Kinders, Robert
APPLICANT: Corey, Michael J.
TITLE OF INVENTION: PAL-18 POLYPEPTIDES, NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING THE SAME AND METHODS FOR SCREENING FOR OR
TITLE OF INVENTION: MODULATING THE SAME
FILE REFERENCE: 130001.406
CURRENT APPLICATION NUMBER: US/09/804.682
NUMBER OF SEQ ID NOS: 174
SCFWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 1064
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 5, 6, 16, 21, 24, 25, 33, 39, 72, 110, 205, 214, 231, 232,
LOCATION: 235, 237, 238, 244, 245, 246, 256, 282, 292, 306, 319,
LOCATION: 321, 323, 330, 334, 340, 349, 354, 355, 372, 376, 378,
LOCATION: 397, 405, 432, 437, 454, 455, 457, 458, 459, 468, 470
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: 485, 487, 488, 494, 496, 499, 511, 524, 527, 552, 557, 562,
LOCATION: 583, 600, 611, 613, 623, 624, 652, 654, 674, 681, 687, 691,
LOCATION: 694, 701, 713, 716, 720, 721, 725, 731, 734, 735, 739, 743,
LOCATION: 744, 781, 782, 785, 789, 799, 803, 821, 823, 847, 852
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: 858, 878, 884, 886, 896, 897, 901, 917, 920, 932, 939, 948,
LOCATION: 957, 961, 965, 981, 991, 993, 1001, 1002, 1005, 1011, 1018,
LOCATION: 1043, 1047, 1049, 1051, 1054, 1056
OTHER INFORMATION: n = A,T,C or G
US-09-804-682-29

Query Match 5.6%; Score 52.6; DB 10; Length 1064;
Best Local Similarity 41.5%; Pred. NC. 0.0047;
Matches 15; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

460 CGTCCGGGGCCGGGTCCGAGCACCCTCCCAAGGGCGCTCGGGCCGGGACCCGCC 519

```

Query Match          5.4%; Score 50.6; DB 10; Length 43058;
Best Local Similarity 46.6%, Pred. NO. 0.02;
Matches 228; Conservative 0; Mismatches 259; Indels 2; Gaps 2
yy      453 GGAGCGCGGTCCCGGGCCGCGTGCACACACTCCACGCGCTCGGCCCGGGCC 512
        ||||| | | | | | | | | | | | | | | | | | | | | |
zb      25300 GAAGCGCGCTGTGTGGAGCTCAGGCCCCCGCCCCGTCCTCAGGCTCTCTC-AGCC 25358

```

Query Match 5.4%; Score 50.6; DB 10; Length 43058;
Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 228; Conservative 0; Mismatches 255; Indels 2; Gaps 2;
QY 513 ACCGCGCTCGTCCGCCGCCGCCCTCTCCGTAGCCAGGAGGAGCCCTGGAGGAA 572
Db 25359 ACCGCGAGCAACAGCTACTTCCCGC-ACITCGACCTGACCCGGGTCCCGCAGTTG 25417
QY 573 GAAGAGGTAGTGGGAGCGGATGAGGGGTGGGACCCCTTGACGTACCAAGAGGA 632
Db 25418 CGCGCGACCGCTCCAAAGTGTGGCGCGGTGGCGACCGGTGAAGAGCATCACAC 25477
QY 633 GTTGGCGGAGTGAAGTGGGTGGGAGAGGTATATAATCGCCGCCCTCGCTGCT 692
Db 25478 ATCGCGGCGCCCTGTCCAAAGTGAAGGTGGGAGAGGTATATAATCGCCGCCCTCGCTGCT 692
QY 693 CTTCATCGAGTCCGCGGAGGCTCGAGCGCGCGGACGACACTCTCTCGGCTCTC 752
Db 25538 GTCAACTTCAAGTGTGCGCGGCGCGGTGCGGCGCGCGGCGCGGCGCGGCGG 25597
QY 753 CCGGCGAGCGCGCGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 812
Db 25598 CCGGCGCGCGCGCGGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 812
QY 813 CCGGCGCGCGCGCGGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 872
Db 25658 CCGGCGCGCGCGCGGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 872
QY 873 GCTCAGCG 941
Db 25778 GGTCCGCGG 25786

RESULT 4

US-09-954-456-529
; Sequence 529, Application US/0954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cano
; FILE REFERENCE: 589290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 529
; LENGTH: 43058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-529

Query Match 5.4%; Score 50.6; DB 10; Length 43058;
Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 228; Conservative 0; Mismatches 255; Indels 2; Gaps 2;
QY 453 GGGGCGCGCTCCCGCGCGCGCGCTCCGTAGCCAGGAGGAGCCCTGGAGGAA 512
Db 25350 GAGGCGCGCGCTCCCGCGCGCGCTCCGTAGCCAGGAGGAGGAGGAGGAGGAGGAA 572
QY 513 ACCGCGCTCGTCCGCCGCCGCCCTCTCCGTAGCCAGGAGGAGGAGGAGGAGGAGGAA 632
Db 25359 ACCGCGAGCAACAGCTACTTCCCGC-ACITCGACCTGACCCGGGTCCCGCAGTTG 25417
QY 573 GAAGAGGTAGTGGGAGCGGATGAGGGGTGGGACCCCTTGACGTACCAAGAGGA 632
Db 25418 CGCGCGACCGCTCCAAAGTGTGGCGCGGTGGCGACCGGTGAAGAGCATCACAC 25477
QY 633 GTTGGCGGAGTGAAGTGGGTGGGAGAGGTATATAATCGCCGCCCTCGCTGCT 692
Db 25478 ATCGCGGCGCCCTGTCCAAAGTGAAGGTGGGAGAGGTATATAATCGCCGCCCTCGCTGCT 692
QY 693 CTTCATCGAGTCCGCGGAGGCTCGAGCGCGCGGACGACACTCTCTCGGCTCTC 752
Db 25538 GTCAACTTCAAGTGTGCGCGGCGCGGTGCGGCGCGCGGCGCGGCGCGGCGG 25597
QY 753 CCGGCGAGCGCGCGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 812
Db 25598 CCGGCGCGCGCGCGGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 812
QY 813 CCGGCGCGCGCGCGGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 872
Db 25658 CCGGCGCGCGCGCGGCTCGAGCGGGTCCGGGGTCCGGGTGCGAGCGCGCGAGCGGCG 872
QY 873 GCTCAGCG 941
Db 25778 GGTCCGCGG 25786

RESULT 5

US-09-880-107-3950
; Sequence 3950, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3950
; LENGTH: 43058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 284721
US-09-880-107-3950

Query Match

Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 228; Conservative 0; Mismatches 255; Indels 2; Gaps 2;
QY 453 GGGGCGCGCTCCCGCGCGCGCGCTCCGTAGCCAGGAGGAGCCCTGGAGGAA 512

Db 25300 GAGCCGCGCTCTGGAGCTCAGCGCCCGCCCGCCCGCTCCAGCTCTTC-ACCC 25358
QY 513 ACCGCGCTCTGTCGCGCCCGCCCGCTCTCGTACGCGAGGAGCGCTGGAGGAA 572
Db 25359 ACCGCGAGCAGACCTACTTCCGC-ACCTCGACCTGACCGCGGCTCCGCGAGTIG 25417
QY 573 GAAGAGGTAGTGGGAGCGGATGAGGGTGGGGACCCCTTGACGTACCAAGGA 632
Db 25418 CGCGCCACCGCTCCAGGTGGTGGCGCGGTGGGAGCGGCTGAGAGCATCAGAC 25477
QY 633 GGTCCCGGGTAGGAAGTGGCTGGGAAAGGTATTAATGCCCCCGCCCTCGCTCT 692
Db 25478 ATCGCGCGCGCTCCAGGTGACGAGCTGCACGCTACATCTCTGGCGTGGACCG 25537
QY 693 CTTATCGAGTTCGCGGAGCTCGGAGCGCGGAGCGGACACCTCTCGGCTCTC 752
Db 25538 GTCAACTTCAAGTTCGCGGCGCGGTGCGGCGCGGCGGCGGCGGCGGCGG 25597
QY 753 CCGCGGAGCGCGGCTCGGAGCGGCTCGGCGGCTCGGCGGCTCGGCGGCGGCGG 812
Db 25598 CGGCGCGCGGCGGCTCGGAGCGGCTCGGCGGCGGCGGCGGCGGCGGCGG 25657
QY 813 CTGCGCGCGGAGGATTACCGCGGAGTGTGCTCTCTGCTGAGCGCGGAGCGGCG 872
Db 25658 CGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 25717
QY 873 GCTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 932
Db 25718 CGGCGCGGAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 25777
QY 933 GGCGCGCGG 941
Db 25778 GGTGCGCGG 25786

RESULT 6

US-09-799-462A-17/C

Sequence 17, Application US/09799462A

Patent No. US20020160970A1

GENERAL INFORMATION:

APPLICANT: Hadlaczky, Gyula

TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF

AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McAuliffe

STREET: 4250 Executive Square, 7th Floor

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/799,462A

FILING DATE: 10-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/835,682

FILING DATE: 10-APR-1997

APPLICATION NUMBER: 08/695,191

FILING DATE: 07-AUG-1996

APPLICATION NUMBER: 08/682,080

FILING DATE: 15-JUL-1996

APPLICATION NUMBER: 08/629,822

FILING DATE: 10-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L

REGISTRATION NUMBER: 33,779
REFERENCE/SEQUENCE NUMBER: 24601-402G

TELECOMMUNICATION INFORMATION:

TELEPHONE: 658-450-8403

TELEFAX: 658-587-5360

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 4299 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

HYPOTHETICAL: N

ANTI-SENSE: NO

FRAGMENT TYPE: <Unknown>

ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 17:

US-09-799-462A-17

Query Match.

Best Local Similarity 5.1%; Score 48.4; DB 9; Length 42999;

Matches 100; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 389 GAGCGCGATGGCAAGAGAGGGGCAAGGCAAGAGGGCGGAGCAAGACCTCTCCGCGCGCG 448

Db 7761 GAGGGGACGGCGCGGGGCGGAGCGGCGCGGCGGCGCGCGCGCGCGCGCGCGCGCG 7702

QY 449 TCCG 508

Db 7701 GGTGCGGGGAGCG 7642

QY 509 GCGCACCG 568

Db 7641 CGGAGCGGGCG 7582

QY 569 GGAAGA 574

Db 7581 GGAGCA 7576

RESULT 7

US-10-125-767-17/C

Sequence 17, Application US/10125767

Patent No. US20020160410A1

GENERAL INFORMATION:

APPLICANT: Hadlaczky, Gyula

TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND

METHODS

FOR PREPARING ARTIFICIAL CHROMOSOMES

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McAuliffe LLP

STREET: 4350 La Jolla Village Drive, 7th Floor

CITY: San Diego

STATE: CA

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/125,767

FILING DATE: 17-Apr-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/724,693

FILING DATE: 28-NOV-2000

APPLICATION NUMBER: 08/835,682

FILING DATE: 10-APR-1997

APPLICATION NUMBER: 08/695,191

FILING DATE: 07-AUG-1996

Matches 144; Conservative 0; Mismatches 125; Indels 2; Gaps 2;

QY 673 CGCCCCCGCTCGGCTCTTCATCGAGTCCGGGAGGCTCGAGCGCGCCAGGCG 732
DB 899 CGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840

QY 733 -GACACTCTCTCGGCTCTCTCCCGGCGGCGGCTCGAGCGGCTCCGCGGCG 751
DB 839 CGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780

QY 792 GGTCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 851
DB 779 GGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720

QY 852 GGTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 910
DB 719 GGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 910

QY 911 GGAATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 941
DB 659 GGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 629

RESULT 10
US-09-894-998-35/c
; Sequence 35, Application US/09894998
; Patent No. US20020090610A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: David C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: HSW-2
US-09-894-998-35

Query Match 4.9%; Score 46.4; DB 10; Length 2481;
Best Local Similarity 55.6%; Pred. No. 0.11;
Matches 89; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

DB 1341 CGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 764

QY 765 GCGCGGCTCGGAGCGGCGGCTCGGCGGCTCGGCGGCGGCGGCGGCGGCGGCGG 824

DB 1281 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1222

QY 825 GATTACCGGGAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 864

DB 1221 CCCCAGGAGTGTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1182

RESULT 11
US-10-023-529-45/c
; Sequence 45, Application US/10023529
; Patent No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING

; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: A. HEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1614);
US-10-023-529-45

Query Match 4.8%; Score 45; DB 12; Length 1614;
Best Local Similarity 46.9%; Pred. No. 0.21;
Matches 206; Conservative 0; Mismatches 230; Indels 3; Gaps 2;

QY 464 CCGCGGCGGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 523

DB 730 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 671

QY 524 TCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 581

DB 670 GCTGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 611

QY 582 AGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 640

DB 610 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 551

QY 641 GGTAGGAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 700

DB 550 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 491

QY 701 AGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 760

DB 490 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 431

QY 761 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 820

DB 430 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 371

QY 821 CGAGGATTAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 880

DB 370 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 311

QY 881 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 899

DB 310 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 292

RESULT 12
US-10-023-523-45/c
; Sequence 45, Application US/10023523
; Patent No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING

```

: TITLE OF INVENTION:  PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
: TITLE OF INVENTION:  ATHEROSCLEROSIS
: FILE REFERENCE:  10797-004001
: CURRENT APPLICATION NUMBER:  US/10/023,523
: CURRENT FILING DATE:  2001-12-17
: PRIOR APPLICATION NUMBER:  US/09/616,289
: PRIOR FILING DATE:  2000-07-14
: PRIOR APPLICATION NUMBER:  US 09/517,849
: PRIOR FILING DATE:  2000-03-02
: PRIOR APPLICATION NUMBER:  US 08/979,608
: PRIOR FILING DATE:  1997-11-26
: PRIOR APPLICATION NUMBER:  US 60/031,930
: PRIOR FILING DATE:  1996-11-27
: PRIOR APPLICATION NUMBER:  US 60/048,547
: PRIOR FILING DATE:  1997-06-03
: NUMBER OF SEQ ID NOS:  53
: SOFTWARE:  FastSeq for Windows Version 4.0
: SEQ ID NO 45
: LENGTH:  1614
: TYPE:  DNA
: ORGANISM:  Homo sapiens
: FEATURE:
: NAME/KEY:  CDS
: LOCATION:  (1)...(1614)
:  S-10-023-523-45

```

Query Match	4.8%;	Score 45;	DB 12;	Length 1614;
Best Local Similarity	46.9%;	Prd. No. 0.21;		
Matches 206;	Conservative	0;	Mismatches 230;	Indels 3; Gaps 2;
QY	464	CCCGGGCCCGGCTCGGACACCTCCCGACGCGGCTCGGCCCGCGGCCACCGCCCTCG	523	
DB	730	CGCGGCGCCACCGCGGCCCTCCGCGCGGCTGTGGCTCGCGCGGCGGCGGCTGTCT	671	
QY	524	TCGGC--CCCGGCCCTCTCCGTACGCCGAGGAACGAGCTGCGAGGAGAGAGAGGT	581	
DB	670	GCTGTGTGGCGCGCGCGGCTGTGGCGCGCGGACGCGCGGCTCCCGGGCGCGA	611	
QY	582	AGTGGGGGCGCGATGAGGGTGGGACCCCTTCACCTACACAGAGAG--GTCCCG	640	
DB	610	CGCGGCGCGGGGCG	551	
QY	641	GGTAGGAGTGGCTGGGAAAGTTATAATGCCCGCCCGCTCGGCTGCTCTTCATCG	700	
DB	550	GCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGCGGCGCGCGCG	491	
QY	701	AGTCTCCCGGAGGCTCGGAGCGGCCACGCGGACACTCTCTCGGCTCTCTCCCGG	760	
DB	490	GGCCAGCGAGGGGCGCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	431	
QY	761	CGCGCGCGGCTCGGAGCGGCTCCGGGCTCGGCTGACGCGGCCAGCGCGCGCTTC	820	
DB	430	CGCGCGCGGCGGACGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCG	371	
QY	821	CGAGGATTACCCGGGGAAGTGGTGTCTCTGTGCTGAGCGCGCGAGCGGGCTCAGG	880	
DB	370	CG	311	
QY	881	CG	999	
DB	310	CGCGCGCGCGCTGGACGGG	292	

```

RESULT 13
US-09-764-877-2718
; Sequence 2718, Application US/09764877
; Patent No. US20030147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17

```

[illegible]

RESULT 14
US-10-023-529-50/c
Sequence 50, Application US/10023529
Patent No. US2602012386A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: La, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/10/023,529
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 09/616,289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: December 8, 2002, 09:24:02 : Search time 37,4297 Seconds
(without alignments)
7710.006 Million cell updates/sec

Title: D64016_COPY_500_1440
Per set score: 941
Sequence: 1 GCCGAAAGACAGGACAC.....GCCGGGTCTTGGCGGGG 941

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Sequences: .441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Data case: Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Res ID	Score	Query Match	Length	DB ID	Description
1	941	100.0	1745	2	US-08-556-424-1
2	941	100.0	1745	4	US-09-263-914-1
3	194.8	20.7	2313	1	US-08-232-538-5
4	194.8	20.7	2313	1	US-09-427-553-1
5	194.8	20.7	2523	4	US-09-051-363-1
6	194.8	20.7	2651	2	US-08-786-164-5
7	55.4	5.9	15378	3	US-08-785-420-1
8	53.2	5.7	1785	1	US-08-729-416C-16
9	51.4	5.5	7218	1	US-08-232-463-14
10	50.6	5.4	801	4	US-09-298-568-3
11	48.6	5.2	4523	4	US-09-473-716-1
12	48.2	5.1	3602	2	US-08-883-795A-33
13	48.2	5.1	3602	4	US-09-018-138-1
14	47.4	5.0	1335	5	PCT-US91-06532-1
15	47.2	5.0	5703	4	US-09-280-590A-36
16	46.8	5.0	1327	4	US-08-483-533-36
17	46.8	5.0	1327	4	US-09-283-471A-36
18	46.6	5.0	7898	4	US-08-984-709A-49
19	46	4.9	46	2	US-08-556-424-4
20	46	4.9	46	4	US-09-263-914-4
21	46	4.9	46	3	US-09-017-631-23
22	45.8	4.9	6453	1	US-08-306-691B-14
23	45.8	4.9	6453	3	US-09-209-668-10
24	45.8	4.9	6453	3	US-09-356-952-8
25	45.8	4.9	9515	1	US-08-920-812-13
26	45.8	4.9	9515	1	US-08-920-827-13
27	45.8	4.9	9515	1	US-08-921-177-13

C 28	45.8	4.9	9515	1	US-08-362-577C-13	Sequence 13, Appl
C 29	45.8	4.9	9515	2	US-08-920-828-13	Sequence 13, Appl
C 30	45.6	4.8	44377	2	US-08-804-227C-7	Sequence 7, Appl
C 31	45.6	4.8	44377	2	US-08-804-198-1	Sequence 1, Appl
C 32	45.4	4.8	4524	2	US-08-845-988-7	Sequence 7, Appl
C 33	45.4	4.8	4524	3	US-09-206-537-7	Sequence 7, Appl
C 34	45.4	4.8	4524	4	US-09-430-834-7	Sequence 7, Appl
C 35	45.4	4.8	4791	4	US-08-949-155-49	Sequence 4, Appl
C 36	45.4	4.8	4791	4	US-09-819-964-49	Sequence 49, Appl
C 37	45.4	4.8	43280	2	US-08-804-227C-1	Sequence 1, Appl
C 38	45.4	4.8	16998	4	US-09-676-610B-24	Sequence 24, Appl
C 39	45.2	4.8	224	3	US-09-017-631-24	Sequence 24, Appl
C 40	45.2	4.8	224	4	US-09-018-138-2	Sequence 2, Appl
C 41	44.8	4.8	1896	4	US-09-343-011B-4	Sequence 4, Appl
C 42	44.8	4.8	1896	4	US-09-343-011B-4	Sequence 4, Appl
C 43	44	4.7	801	2	US-08-770-379-16	Sequence 16, Appl
C 44	44	4.7	801	4	US-08-757-669A-16	Sequence 16, Appl
C 45	44	4.7	801	4	US-09-230-371A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-556-424-1
Sequence 1, Application US/08556424
Patent No. 5916763
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Morishita, Kaoru
TITLE OF INVENTION: No. 5916763el Promoter for VEGF Receptor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,424
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0635000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1155
OTHER INFORMATION: /note= "Nucleotides numbered 1
OTHER INFORMATION: through 1195 correspond to -1195 through -1 from
OTHER INFORMATION: Figure 2."
US-08-556-424-1

Query Watch 100.0% Score 941, DB 2; Length 1745;
Best Local Similarity 100.0%; Pred. No. 1.6e-191;
Matches 941; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCGAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 60
Dd 500 GCGAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 559
QY 61 GTCCAAAGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 120
Dd 560 GTCCAAAGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 619
QY 121 CTCGGTGCCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 180
Dd 620 CTCGGTGCCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 679
QY 181 GCGGAAAGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 240
Dd 680 GCGGAAAGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 739
QY 241 ACCGTGCGGAGGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 300
Dd 740 ACCGTGCGGAGGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 799
QY 301 GCGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 360
Dd 800 GCGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 859
QY 361 GCGTGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 420
Dd 860 GCGTGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 919
QY 421 AGAGGCGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 480
Dd 920 AGAGGCGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 579
QY 481 AGACCTCCACGCGGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 540
Dd 980 AGACCTCCACGCGGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1039
QY 541 CCGTGGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 600
Dd 1040 CCGTGGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1099
QY 601 GCGTGGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 660
Dd 1100 GCGTGGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1159
QY 661 AGGTATTAATATGCGGCGGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 720
Dd 1160 AGGTATTAATATGCGGCGGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1219
QY 721 GCGGCGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 780
Dd 1220 GCGGCGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1279
QY 781 CTCGGGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 840
Dd 1280 CTCGGGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1339
QY 841 GGTTCCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 900
Dd 1340 GGTTCCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1399
QY 901 ACAGAGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 941
Dd 1400 ACAGAGGAGGAGCTTCAGCTCTTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 1440

RESULT 2
US-09-263-914-1
; Sequence 1, Application US/09263914
; Patent No. 6245512
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Morishita, Keoru

```

Db 860 GGGTGGAGGAGTCTGCAAGGATTTCCTAGCCGATGGGAGGAGGAGGCGCAAGGCA 919
QY 421 AGAGGCGCGGAGCAACACCTGAACTTCCCGGGCGCGCTCCCGGGCGCGGTCGCC 480
Db 920 AGAGGCGCGGAGCAACACCTGAACTTCCCGGGCGCGCTCCCGGGCGCGGTCGCC 979
QY 481 AGCACTTCCCGGAGGCGCTCGGCGCGCGGCGCGGCGCGGCGCGGCGCGGCGCTCT 540
Db 980 AGCACTTCCCGGAGGCGCTCGGCGCGCGGCGCGGCGCGGCGCGGCGCGGCGCTCT 1039
QY 541 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 1040 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
QY 601 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 1100 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1159
QY 661 AAGTTATATAATGCG 720
Db 1160 AAGTTATATAATGCG 1219
QY 721 GCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 1220 GCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1279
QY 781 CTCGGGCGGTCGGGTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Db 1280 CTCGGGCGGTCGGGTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1339
QY 841 GGTGTCTCTGCTGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 1340 GGTGTCTCTGCTGGAGCGCGGAGAGCGGCGGCTCAGGCGCGCGCGCGCGCGCGCGCGCG 1399
QY 901 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 941
Db 1400 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

RESULT 3

US-08-232-538-5
; Sequence 5, Application US/08232538
; Patent No. 5712380
; GENERAL INFORMATION:
; APPLICANT: THOMAS, KENNETH A.
; APPLICANT: KENDALL, RICHARD L.
; TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
; GROWTH FACTOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000 126 E Lincoln Avenue
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,538
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen, John W.III
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 188881A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2313 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-232-538-5

Query Match

20.7%; Score 194.8; DB 1; Length 2313;
Best Local Similarity 98.1%; Pred. No. 3.7e-33;
Matches 208; Conservative 0; Mismatches 2; Indels 1;

QY 730 GCGGACACTCTCTCGGCTCTCTCCCGGAGCGCGGCGGCGGCTCGGAGCGGGCTCCGGGGC 789
Db 1 GCGGACACTCTCTCGGCTCTCTCCCGGAGCGCGGCGGCGGCTCGGAGCGGGCTCCGGGGC 60
QY 790 TCGGGTGCAGCGCGCGCGCGCGCGCTCGGCGGAGGATTACCGGGGAAGTGGTGTCTC 849
Db 61 TCGGGTGCAGCGCGCGCGCGCGG--GCCIGCGCGAGGATTACCGGGGAAGTGGTGTCTC 118
QY 850 CTGCTGGAGCGCGGAGCGGCGCTCAGGGCGCGGGCGCGCGCGGCGGCGGCGGCGGCGG 909
Db 119 CTGCTGGAGCGCGGAGCGGCGCTCAGGGCGCGGGCGCGCTCAGGGCGCGGGCGCGGCGG 178
QY 910 CGGACTCTGCGCGCGCGGCTCTTGGCGCGGG 941
Db 179 CGGACTCTGCGCGCGCGGCTCTTGGCGCGGG 210

RESULT 4

US-09-427-353-1
; Sequence 1, Application US/09427353
; Patent No. 6375929
; GENERAL INFORMATION:
; APPLICANT: THOMAS, KENNETH A.
; APPLICANT: GOLDMAN, COREY K.
; APPLICANT: KENDALL, RICHARD L.
; APPLICANT: BETT, ANDREW J.
; APPLICANT: HUCKLE, WILLIAM R.
; TITLE OF INVENTION: GENE THERAPY FOR INHIBITION OF
; ANGIOGENESIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,353
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19810YCA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2313 base pairs
; TYPE: nucleic acid


```

; LENGTH: 15378 base pairs
; TYPE: nucleic acid

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944

RESULT 6
 US-08-729-416C-16
 : Sequence 16, Application US/08729416C
 : Patent No. 6013767
 : GENERAL INFORMATION:
 : APPLICANT: NAKAMURA, TAKESHI
 : TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GEN
 : TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
 : NUMBER OF SEQUENCES: 22
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
 : STREET: 1100 NEW YORK AVENUE, N.W.
 : CITY: WASHINGTON
 : STATE: D.C.
 : COUNTRY: USA
 : ZIP: 20005-3918
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentLib Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/729-416C
 : FILING DATE: 11-OCT-1996
 : CLASSIFICATION: 536
 : ATTORNEY/AGENCY INFORMATION:
 : NAME: PERRY, GLENN J.
 : REGISTRATION NUMBER: 28458
 : REFERENCE/DOCKET NUMBER: 7898/225948
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202-861-3000
 : TELEFAX: 202-822-0944
 : TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1785 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double strand
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 US-08-729-415C-16

Query Match 5.7%; Score 53.2; DB 3; Length 1785;
 Best Local Similarity 51.7%; Pred. No. 0.0041;
 Matches 121; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
 QY 674 GCCCGCCGCTCTTCATCGAGTCCGCGGAGCTCGAGCGGCCAGGCGG 733
 Db 224 CTCGCGCTCGTGGCTCTCTCGCGCGCCGCGGAGCGGAGCGGCTCTGG 263
 QY 734 ACACTCTCTCTCGGCTCTCTCCCGGAGCGGCGGCGTCTCGAGCGGCTCCGGGCTCGG 793
 Db 284 CAGCTGCGAGCGCCAGCTCGGCGGCGGCGGAGTGGCAGCGACCCCGGTCC 343
 QY 794 GTGACCGCGAGCGGCGCTCGGCGGAGGATTACCGGGAGTGTCTCTCTGG 853
 Db 344 GTGCGCGCGCGGCGGAGTGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 403
 QY 854 GTGAGCGCGGAGCGGCGCTCAGCGCGGCGGCGGCGGCGGCGGCGGCGGAG 907
 Db 404 CAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 457

RESULT 9
 US-08-232-163-14/c
 Sequence 14, Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHEIFLINGER, E.
 APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)833-4109
 TELEX: 895149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: pTZ9pl-Fls
 US-08-232-463-14

Query Match 5.5%; Score 51.4; DB 1; Length 7218;
 Best Local Similarity 3.8%; Pred. No. 0.011;
 Matches 13; Conservative 197; Mismatches 133; Indels 0; Gaps 0;
 QY 346 AAGCTGAGTACCGGCTCGAGGAGTCTGCAAGATTCTTCTGACGCGATGGCAGGA 405
 Db 1391 KRR 1332
 QY 406 GGAGGGCAAGGCGGCGGAGCGGCGGAGCAAGACCTTGACCTGCGGGCGCGGCTCC 465
 Db 1331 KRR 1272
 QY 466 CGGCGCCCGCTCGCCACCACTCCCAACCGCGCTCGGCGGCGGCGGCGGCGGCTC 525
 Db 1271 KRR 1212
 QY 526 GCGCGCCCGCCCTCTCGTAGCGAGGAGGAGGAGGCTGGAGGAGAGAGAGGTAGT 585
 Db 1211 KRR 1152
 QY 586 GGAGGAGCGGATGAGGGTGGGGGACCCCTTACGTACCAAGAGAGAGGTGCGGGTAG 645
 Db 1151 KRR 1092
 QY 646 GAAGTGGCGCGGAGAGGTATAATCGCCCGCGGCGGCGGCGGCGGCGGCGG 688
 Db 1091 KRR 1049

RESULT 10
 US-09-298-568-3/c
 Sequence 3, Application US/09298568
 Patent No. 6322752
 GENERAL INFORMATION:
 APPLICANT: Kieff, Elliott D.
 APPLICANT: Ballestas, Mary E.
 APPLICANT: Kaye, Kenneth M.
 TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
 TITLE OF INVENTION: VIRUS DNA TO MEDIANE EFFICIENT EPISOME PERSISTENCE
 FILE REFERENCE: 16412-10001R
 CURRENT APPLICATION NUMBER: US/09/298,568
 CURRENT FILING DATE: 1999-04-21
 EARLIER APPLICATION NUMBER: US 60/109,422
 EARLIER FILING DATE: 1998-11-19
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO: 3
 LENGTH: 801
 TYPE: DNA
 ORGANISM: Kaposi's sarcoma-associated herpesvirus
 US-09-298-568-3
 Query Match 5.4%; Score 50.6; DB 4; Length 801;
 Best Local Similarity 44.2%; Pred. No. 0.013;
 Matches 302; Conservative 0; Mismatches 374; Indels 7; Gaps 2;
 QY 216 CCTCAGCTCTCTGTCGCAAGACACCGTTCGGGAGGCGGCGGCGGCGGCGGCGG 277
 Db 666 CCGGGGTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 625
 QY 278 ACTTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 337
 Db 628 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 569
 QY 338 CGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 397
 Db 568 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 509

Db	442	CGCCGCGCAGCGCGGGCGCGCCAGACGAGCGGGCGAGCGACCGGGCGCCCTCGGGGCG	541
QY	753	CGCGCGATCGAGCGGGCTCCGGGCTCGGTGCAGCGGCACCGGGCGCTGCGGGCG	822
Db	542	CGCGCGGGCTCGCGCGCGCGCTGCTCGCGGGGGGACGAGAGTGGCGCTCGCT	601
QY	823	AGGATTACCGGGGAAGTGGTGTCTCCCTGGAGCCCGAGACGGCGCTCAGGGCG	882
Db	602	CGGTGAGGTGGGTCTGAGGACCGCGCGGCGAAGGGGCGCGGGCGACGAGCTGGAGG	661
QY	893	CGGGCGCGCGGGCGAGCAACGAGACGAGCTCTGCGGGCGG	925
Db	662	CCGCGCGCTCTGAGGGCGGGCCAGGGGTCCGGGATGGCGGCGAG	704

RESULT 12

Patent No. 5985607
GENERAL INFORMATION:

APPLICANT: Awang, Gregor

	TITLE OF INVENTION:	RECORD
:	TITLE OF INVENTION:	VECTOR
:	TITLE OF INVENTION:	VECTOR

CORRESPONDENCE ADDRESS:

STREET: 40 King Street

STATE: Ontario

; ZIP: M5H 3Y2
; CONVERTED FROM 3000

MEDIUM TYPE: Floppy diskette

OPERATING SYSTEM: PC-DOS
SOFTWARE: Patent In Release

: CURRENT APPLICATION DATA: USA
 : APPLICATION NUMBER:

CLASSIFICATION: 435

NAME: Gravelle, Michelle

REFERENCE/DOCKET NUMBER:

TELEPHONE: (416) 364-7331

; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 3602 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear

```

MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:

ORGANISM: HOMO SAPIENS
FEATURE:

LOCATION: join(625..637)

FEATURE:

```

; LOCATION: join(625..637

```

US-08-883-795A-33

Query Match	5.18

Matches 122; Conservative

QY 006 TATAAATCGCCCCCGCCCTCG

Db 644 TACTACCGTGACACCCCTATCTCCGGCTGCGGGTCCCTCAGACGACCTGGGGCGG 585
QY 726 CCAGCGGACACTCTCTCGGCTCTCCCGGAGCGCGGCTCGGAGCGGCTCCG 785
Db 584 CCGGTGACACCGCGGCGCTCTCCCGGAGCTCGGGTCCGAGCGGCGGCGG 525
QY 786 GGGCTGCGGTGACCGCGGCGGCTCTCCCGGAGCTCCGAGGAGTACCGGGGAGTGG 845
Db 524 CCGGCTAGAGGAGCGGCGCTCTCCAGGGCGGCGGTCTCGGAGCAGCGGCGCGG 465
QY 846 TCTCTGCTGAGCGCGGCGGCGGCTCTCCAGCGGCGGCGGCGGCGGCGGCGGAG 905
Db 464 TGGCGCGGCTCGGCTCTCCCGGAGGAGCGGCTCTCCGCGGCGGCGGCGGCGG 405
QY 906 AGGAC 910
Db 404 GTGGC 400

RESULT 13
09-018-138-1/c
Sequence 1, Application US/09018138
Patent No. 6187564
GENERAL INFORMATION:
APPLICANT: Sytkowski, Arthur J.
TITLE OF INVENTION: Production and Use of Recombinant
FILE REFERENCE: BH97-05A
CURRENT APPLICATION NUMBER: US/09/018.138
CURRENT FILING DATE: 1998-02-03
EARLIER APPLICATION NUMBER: 08/890,929
EARLIER FILING DATE: 1997-07-10
NUMBER OF SEQ ID NOS: 24
SOFTWARE: fastseq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3602
TYPE: DNA
ORGANISM: Human
US-09-018-138-1

Query Match 5.1%; Score 48.2; DB 4; Length 3602;
Best Local Similarity 49.8%; Pred. No. 0.05;
Matches 122; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
QY 666 TATAATCGCCCGCCCTCGGCTCTCTTCATCGAGGTCGCGGAGGCTCGGAGCGG 725
644 TACTACCGGTGACACCCCTATCTCCGGCTGCGGGTCCCTCAGACGACCTGGGCGG 585
726 CCAGCGGACACTCTCTCGGCTCTCCCGGAGCGGCGGCTCGGAGCGGCTCCG 785
584 CCGGTGACACCGCGGCGGCTCTCCCGGAGCTCCGCGGCTCGGAGCGGCGGCGG 525
QY 786 GGGCTGCGGTGACCGCGGCGGCTCTCCCGGAGGATTAACCGGGGAGTGG 845
Db 524 CCGGCTAGAGGAGCGGCGCTCTCCAGGGCGGCGGTCTCGGAGCAGCGGCGCGG 465
QY 846 TCTCTGCTGAGCGCGGCGGCTCTCCAGCGGCGGCGGCGGCGGCGGCGGAG 905
Db 464 TGGCGCGGCTCGGCTCTCCCGGAGGAGCGGCTCTCCGCGGCGGCGGCGG 405
QY 906 AGGAC 910
Db 404 GTGGC 400

RESULT 14
PCT-US91-06532-1/c
Sequence 1, Application PCT/US9106532
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Hicknell
STREET: Two First National Plaza Suite 2100
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION LAIR:
APPLICATION NUMBER: PCT/US91/06532
FILING DATE: 19910910
CLASSIFICATION: 42
ATTORNEY/AGENT INFORMATION:
NAME: Gruber, Lewis S.
REGISTRATION NUMBER: 30,060
REFERENCE/DOCKET NUMBER: 27373/5235
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1335 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US91-06532-1

Query Match 5.0%; Score 47.4; DB 5; Length 1335;
Best Local Similarity 46.0%; Pred. No. 0.067;
Matches 271; Conservative 0; Mismatches 311; Indels 7; Gaps 3;
QY 357 ACCCGGCTGAGGAGTCTCGAGGATTCCTGAGGCGGATGGGACGAGAGGGGCAAG 416
Db 946 ACCGGCTTGGTGGGAGCTCCCGGGAGCTCCGGGAGAGACCCAGGCGGCTCGGGTGTA 527
QY 417 GCGAAGAGGCGGCGGAGACCTCTGACCTGCGGGGCGGCGCTCCCGGGCGCGG 476
Db 926 GTTATAGACGAGTTCGCGGGCGGCTCCCGGGCGGAGGCGGCGGCGGCGGCGG 867
QY 477 GCGGAGGACCTCCCGAGCGGCTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 536
Db 866 CCGAGGACGCGGCGGAGGAGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGG 808
QY 537 CTCCTCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 596
Db 807 GTGCGGCGGCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 748
QY 557 TACGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 556
Db 747 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 688
QY 657 GGGAAAGGATTAATATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 716
Db 687 CCGGCGGCTCGGCGGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 628
QY 717 CGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 776
Db 627 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 568
QY 777 CCGGCTCCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 836
Db 567 CCGGCGGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 509
QY 837 AAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 896
Db 508 AGGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 449

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:50:37 ; Search time 189.314 Seconds
(without alignments)
11193.756 Million cell updates/sec

Title: D64016_COPY_500_1440
Perfect score: 941
Sequence: 1 GCGGAAAGACGACGAC.....GCCGGTCITTCGCCGGG 941

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

tal number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
N_Geneseq_101002:**			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:**		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:**		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:**		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:**		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:**		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:**		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:**		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:**		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:**		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:**		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:**		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:**		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:**		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:**		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:**		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:**		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:**		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:**		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:**		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:**		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:**		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:**		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:**		
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:**		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	941	100.0	1745	18	Flt-1 VEGF receptor
2	667	70.9	840	17	Promoter of human
3	613	65.1	6316	22	Tumour suppressor
4	613	65.1	6316	24	Chemically treated
5	613	65.1	6316	24	Human gene regulat
6	613	65.1	6316	24	Signal transductio
7	613	65.1	6321	24	Human angiogenesis
8	510.8	54.3	6316	22	Tumour suppressor
9	510.8	54.3	6316	24	Chemically treated

C 10	510.8	54.3	6316	24	AA561171	Human gene regulat
C 11	510.8	54.3	6316	24	ABK31261	Signal transductio
C 12	510.8	54.3	6321	24	ABQ67064	Human angiogenesis
C 13	194.8	20.7	2313	19	AAV19347	Human soluble VEGF
C 14	194.8	20.7	2313	19	AAV09330	Soluble VEGF recep
C 15	194.8	20.7	2523	18	AAV01457	Human VEGF recepto
C 16	194.8	20.7	2523	18	AA162101	CDNA encoding amin
C 17	194.8	20.7	2651	15	AAQ74268	SVGF-R1 gene. Ho
C 18	194.8	20.7	2651	20	AAQ04428	Human soluble vasc
C 19	194.8	20.7	7680	23	AA570286	DNA encoding novel
C 20	130.6	13.9	3009	22	AA591131	Flt-1extraFASm/Cy
C 21	69	7.3	11355	20	AA553491	Human adenosine A1
C 22	64.2	6.6	11355	20	AA553461	Human adenosine A1
C 23	63.6	6.8	1337	20	AA17263	Human gene express
C 24	59	6.3	1127	21	AAA02477	Human colon cancer
C 25	57.8	6.1	1059	21	AAA01833	Human colon cancer
C 26	56.2	6.0	154746	24	AAQ25519	Human herpesvirus
C 27	56.2	6.0	154746	24	AAQ25519	Human herpesvirus
C 28	55.6	5.9	600	24	AB052496	Oligonucleotide fo
C 29	55.6	5.9	600	24	AB052497	Oligonucleotide fo
C 30	55.4	5.9	13377	13	AAQ25975	M4 mutant porcine
C 31	55	5.8	1218	21	AAA02488	Human colon cancer
C 32	54.8	5.8	1000	21	AAA02484	Human colon cancer
C 33	53.6	5.7	989	21	AAA02539	Human colon cancer
C 34	53.2	5.7	1785	18	AA166990	Rat brain-specific
C 35	51.6	5.5	754	24	AB040800	Oligonucleotide fo
C 36	51.6	5.5	754	24	AB040801	Oligonucleotide fo
C 37	51	5.4	629	24	AB043450	Oligonucleotide fo
C 38	51	5.4	629	24	AB043451	Oligonucleotide fo
C 39	50.8	5.4	3465	22	AA54723	Nucleotide sequenc
C 40	50.6	5.4	801	21	AAA30291	Rhadinovirus cis-
C 41	50.6	5.4	801	24	AB493488	Kaposi's sarcoma-a
C 42	50.6	5.4	43058	24	ABN97455	Gene #3953 used to
C 43	50.6	5.4	43058	24	AB164982	Lung cancer relate
C 44	50.6	5.4	43058	24	AB165219	Lung cancer relate
C 45	50.4	5.4	15672	12	AAQ10613	Rianodin receptor

ALIGNMENTS

RESULT 1
AA169249
ID AA169249 standard, DNA; 1745 BP.

XX AA169249;

XX 11-JAN-1998 (first entry)

DE Flt-1 VEGF receptor gene promoter region.

DE Flt-1: VEGF receptor; vascular endothelial growth factor; human;
KW vascular permeability factor; fms-like receptor tyrosine kinase;
KW hyperextension; thrombosis; atherosclerosis; restenosis;
KW inflammatory disease; haemophilia; wound healing;
KW diabetic retinopathy; rheumatoid arthritis; blood cell trafficking;
KW tumour; angiogenesis; gene therapy; transgenic animal; ss.

XX Homo sapiens.

Key	Location/Qualifiers
EH Promoter	447...1479
FI	/tag= a
FI	/note= "(Claim 4)"
FI	1051...1055
FI	/tag= b
FI	/label= ets
FI	/note= "E26 transformation specific sequence"
FI	1068...1071
FI	/tag= c
FI	/label= ets
FI	/note= "E26 transformation specific sequence"
FI	1098...1103
FI	misc_feature


```
QY 776 GCGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 835
Db 61 GCGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 120
QY 836 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGCGGCGGCGG 895
Db 121 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGCGGCGGCGG 180
QY 896 GCGAAGCAGAGGAGCGGACTCTGGCGCGCGGCTTGGCGCGGG 941
Db 181 GCGAAGCAGAGGAGCGGACTCTGGCGCGCGGCTTGGCGCGGG 226

RESULT 6
AL543052 972 bp mRNA linear EST 16-FEB-2001
LOCUS AL543052 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1002Y109.5
DEFINITION prime, mRNA sequence.
ACCESSION AL543052
VERSION AL543052.1 GI:12875530
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 972)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 972
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1002Y109"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life technologies. Contact : Feng Liang Life technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 259 a 248 c 266 g 194 t 5 others
ORIGIN
Query Match 23.7% Score 223.4; DB 9; Length 972;
Best Local Similarity 59.1%; Pred. No. 3.1e-33;
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 716 TCGGAGCGGGCGGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 775
Db 1 TCGGAGCGGGCGGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 50
QY 776 GCGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 835
Db 61 GCGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 120
QY 836 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGCGGCGGCGG 895
Db 121 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGCGGCGGCGG 180
QY 896 GCGAAGCAGAGGAGCGGACTCTGGCGCGCGGCTTGGCGCGGG 941
Db 181 GCGAAGCAGAGGAGCGGACTCTGGCGCGCGGCTTGGCGCGGG 226

RESULT 8
AL136704 662 bp mRNA linear EST 02-AUG-2002
LOCUS AL136704 PLACE1 Homo sapiens cDNA clone PLACE1004901.5, mRNA
DEFINITION sequence.
ACCESSION AL136704
VERSION AL136704.1 GI:10997243
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
RESULT 7
AL547815 1002 bp mRNA linear EST 16-FEB-2001
LOCUS AL547815 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1033YD17.5
DEFINITION prime, mRNA sequence.
ACCESSION AL547815
VERSION AL547815.1 GI:12882232
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1002)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 1002
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1033YD17"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life technologies. Contact : Feng Liang Life technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 251 a 240 c 281 g 229 t 1 others
ORIGIN
Query Match 23.6% Score 221.8; DB 9; Length 1002;
Best Local Similarity 99.1%; Pred. No. 6.3e-33;
Matches 223; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 717 CGAGCGCGCGGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 776
Db 1 CGAGCGCGCGGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 777 GCGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 836
Db 61 GCGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 120
QY 837 AAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGGCGGCGGCGG 896
Db 121 AAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGGCGGCGGCGG 180
QY 897 GCGAAGCAGAGGAGCGGACTCTGGCGCGCGGCTTGGCGCGGG 941
Db 181 GCGAAGCAGAGGAGCGGACTCTGGCGCGCGGCTTGGCGCGGG 225

RESULT 8
AL136704 662 bp mRNA linear EST 02-AUG-2002
LOCUS AL136704 PLACE1 Homo sapiens cDNA clone PLACE1004901.5, mRNA
DEFINITION sequence.
ACCESSION AL136704
VERSION AL136704.1 GI:10997243
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE *
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 562)
Ota, T., Nishikawa, I., Suzuki, Y., Ishii, S., Saito, K., Kawa, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, J., Sugano, S. and Isogai, T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5' & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute
Location/Qualifiers
1. 562

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PLACE104901"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: pME18SF13"

BASE COUNT 153 a 175 c 199 g 121 t 14 others

ORIGIN

Query Match 22.9%; Score 215.4; DB 9; Length 662;
Best Local Similarity 97.7%; Pred. No. 1.1e-31;
Matches 216; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 721 GCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGAGCGGG 780
Db 1 GCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGAGCGGG 60
QY 781 CTCGGGCTCGGTCGACGCGGACGCGGCGCTCGGCGGAGGATACCCGGGAGI 840
Db 61 CTCGGGCTCGGTCGACGCGGACGCGGCGCTCGGCGGAGGATACCCGGGAGI 120
QY 841 GGTCTCTCTCGGTCGAGCGGACGCGGCGCTCAGGCGGCGGCGGCGGCGGCGA 900
Db 121 GGTCTCTCTCGGTCGAGCGGACGCGGCGCTCAGGCGGCGGCGGCGGCGGCGA 180
QY 901 ACGAGAGCGGACACTCTCGGCGGCGGCGGCTTTTGCCCGCGGG 941
Db 181 ACGAGAGCGGACACTCTCGGCGGCGGCGGCTTTTGCCCGCGGG 221

RESULT 9

AL551816 1025 bp mRNA linear EST 16-FEB-2001
DEFINITION AL551816 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1051918.5
prime, mRNA sequence.

ACCESSION AL551816
VERSION AL551816.1 GI:12890126
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 1025)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 1025
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="CS0D1051918"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com

BASE COUNT 282 a 255 c 268 g 207 t 13 others

ORIGIN

Query Match 22.6%; Score 212.2; DB 9; Length 1025;
Best Local Similarity 97.8%; Pred. No. 4.3e-31;
Matches 222; Conservative 3; Mismatches 1; Indels 1; Gaps 1;
QY 715 TCGGAGCGCCGAGCGGACACTCTCTCGGCTCTCTCCCGGACGCGGCGGCTCGGA 775
Db 1 TCGGAGCGCCGAGCGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCTCGGA 60
QY 775 CCGGCTCCCGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 834
Db 61 CCGGCTCCCGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 835 GGAAGTGTGTCTCTCTCGGCTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 894
Db 121 GGAAGTGTGTCTCTCTCGGCTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 895 CCGGCAACGAGAGCGGACACTCTCGGCGGCGGCGGCTTTTGCCCGCGGG 941
Db 181 CCGGCAACGAGAGCGGACACTCTCGGCGGCGGCGGCTTTTGCCCGCGGG 227

RESULT 10

AL548886 538 bp mRNA linear EST 16-FEB-2001
DEFINITION AL548886 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1043YK17.5
prime, mRNA sequence.

ACCESSION AL548886
VERSION AL548886.1 GI:12884334
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 538)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 538
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1043YK17"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : filanglifetech.com URL :

http://fulllength.invitrogen.com

BASE COUNT 148 a 176 c 202 g 108 t 4 others
 ORIGIN

Query Match 22.5%; Score 211.4; DB 9; Length 538;
 Best Local Similarity 99.1%; Pred. No. 5.2e-31;
 Matches 223; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 718 GGAGCGCGCAGCGGACCTCTCTCGGCTCTCCCGGAGCGGCGGCGCTCGGAGC 777

Db 1 GGAGCGCGCAGCGGACCTCTCTCGGCTCTCCCGGAGCGGCGGCGCTCGGAGC 60

Oy 778 GGCTCCGGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 836

Db 61 GGCTCCGGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120

Oy 837 AAGTGTCTCTCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 896

Db 121 AAGTGTCTCTCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180

897 GCGACGAGGAGCGGACTCTGCGGCGGCGGCTCTTGGCGCGGG 941

Db 181 GCGACGAGGAGCGGACTCTGCGGCGGCGGCTCTTGGCGCGGG 225

RESULT 11

AL546515

LOCUS

AL546515 L11_NFL006_PL2 Homo sapiens cDNA clone CS0D1030YL03 5

DEFINITION

prime, mRNA sequence.

ACCESSION

AL546515

VERSION

AL546515.1 GI:12879706

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 860)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

CONTACT: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..860

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1030YL03"

/clone_lib="L11_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with NotI and

cloned into the NotI and EcoRV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life technologies. Contact : Feng Liang life technologies.

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 510 8371

Email : filanglifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 220 a 220 c 224 g 171 t 25 others

ORIGIN

Query Match

Best Local Similarity 22.0%; Score 207.4; DB 9; Length 850;

Matches 208; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 730 GCGACACTCTCTCGGCTCTCCCGGAGCGGCGGCTCGGAGCGGCTCCGGGCG 789

Db 1 GCGACACTCTCTCGGCTCTCCCGGAGCGGCGGCTCGGAGCGGCTCCGGGCG 60

Oy 790 TCGGTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 849

Db 51 TCGGTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120

Oy 850 CTGCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 909

Db 121 CTGCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180

Oy 910 CGGACTCTGCGGCGGCGGCGGCTCTTGGCGCGGG 941

Db 181 CGGACTCTGCGGCGGCGGCGGCTCTTGGCGCGGG 212

RESULT 12

BO271447

LOCUS

ik14a07.yl HR85 islet Homo sapiens cDNA clone IMAGE: 5780553 5'

similar to TR-Q12954 Q12954 SOLUBLE VASCULAR ENDOTHELIAL CELL

GROWTH FACTOR RECEPTOR. ; mRNA sequence.

ACCESSION BO271447

VERSION BO271447.1 GI:20496513

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 584)

Wellton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blustein, A.,

Schmitt, A., Theisinger, B., Ritter, E., Ronko, I., Bennett, J., Cardenas

M., Gibbons, M., McCann, R., Cole, R., Isagareishvili, R., Williams, I.,

Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other_ESIS: ik14a07.xl

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@bioph.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Dr. Hiroshi Inoue

(hinoue@im.wustl.edu)

Seq primer: -40bp from Gibco

High quality sequence stop: 476.

Location/Qualifiers

1..584

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE: 5780553"

/clone_lib="HR85 islet"

/tissue_type="Purified pancreatic islet"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:

NotI; Site_2: XhoI; cDNA made by oligo-dT priming.

Size selected on agarose gel. Average insert size -1kb. 5'

XhoI site was destroyed after directional cloning.

Amplified once. Contact information: Hiroshi Inoue, MD,

Metabolism Div. (Alan Permutt Lab), Washington University

School of Medicine, Box 8127, 660 South Euclid Ave., St.

Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:

314-362-1916, Fax: 314-747-2692."

BASE COUNT 126 a 163 c 189 g 106 t

ORIGIN

Query Match

Best Local Similarity 22.0%; Score 206.6; DB 14; Length 584;

Matches 208; Conservative 98.1%; Pred. No. 5.2e-30;

```
Matches 209: Conservative 0: Mismatches 4: Indels 0: Gaps 0:
QY 723 GCGGACACTCTCTCGCTCTCTCCCGCAGCGCGCGCTCGAGCGGCTCCCGGG 788
      |||
Db 1 GCGGACACTCTCTCGCTCTCTCCCGCAGCGCGCGCTCGAGCGGCTCCCGGG 60
QY 789 CTCGGGTGAGCGCCAGCGCGCTCTCGCGGAGGATACCGCGGAGTGTGTCT 848
      |||
Db 61 CTCGGGTGAGCGCGCGCGCTCTCGCGGAGGATACCGCGGAGTGTGTCT 120
QY 849 CTGCTGAGCGCGCGCGCTCTCGCGGAGGATACCGCGGAGTGTGTCT 908
      |||
Db 121 CTGCTGAGCGCGCGCGCTCTCGCGGAGGATACCGCGGAGTGTGTCT 180
QY 909 ACGGACTCTCGCGCGCGCTCTCTTTGGCGCGGG 941
      |||
Db 181 ACGGACTCTCGCGCGCGCTCTCTTTGGCGCGGG 213

SULT 13
AL543703
DEFINITION AL543703 LTI_NFL006_PL2 1030 bp mRNA linear EST 16-FEB-2001
prime, mRNA sequence.
ACCESSION AL543703
VERSION AL543703.1 GI:12876182
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1030)
Li, W.B., Gruber, C., Jessee, J. and Pollayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 1030
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1005YF20"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pcwvspor1 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcwvspor1 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies.
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 276 a 267 c 272 g 209 t 5 others
ORIGIN
Query Match 21.9%; Score 206.4; DB 9: Length 1030;
Best Local Similarity 99.1%; Pred. No. 5.6e-30;
Matches 218: Conservative 0: Mismatches 1: Indels 1: Gaps 1:
QY 723 GCGGACACTCTCTCGCTCTCTCCCGCAGCGCGCGCTCGAGCGGCTCCCGGG 782
      |||
Db 1 GCGGACACTCTCTCGCTCTCTCCCGCAGCGCGCGCTCGAGCGGCTCCCGGG 60
QY 783 CTCGGGTGAGCGCGCGCTCTCGCGGAGGATACCGCGGAGTGTGTCT 841
      |||
Db 61 CTCGGGTGAGCGCGCGCTCTCGCGGAGGATACCGCGGAGTGTGTCT 120
QY 842 GTTGTCTCTCTCGCTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 901
      |||
```

```
Db 121 GTTGTCTCTCTCGCTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 180
QY 942 CAGAGCGGACACTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 941
      |||
Db 121 CAGAGCGGACACTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 220

RESULT 14
AL549158
LOCUS AL549158 LTI_NFL006_PL2 740 bp mRNA linear EST 16-FEB-2001
DEFINITION AL549158 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1051YF21 5
prime, mRNA sequence.
ACCESSION AL549158
VERSION AL549158.1 GI:12884867
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 740)
Li, W.B., Gruber, C., Jessee, J. and Pollayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1051YF21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pcwvspor1 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcwvspor1 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies.
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 177 a 195 c 228 g 138 t 1 others
ORIGIN
Query Match 21.7%; Score 204.4; DB 9: Length 740;
Best Local Similarity 99.1%; Pred. No. 1.4e-29;
Matches 216: Conservative 0: Mismatches 1: Indels 1: Gaps 1:
QY 725 GCGGACACTCTCTCTCGCTCTCTCCCGCAGCGCGCGCTCGAGCGGCTCC 784
      |||
Db 1 GCGGACACTCTCTCTCTCGCTCTCTCCCGCAGCGCGCGCTCGAGCGGCTCC 60
QY 785 GCGGCTCGGCTGAGCGCGCGCAGCGCGCGCTCGCGGAGTGTGTCT 843
      |||
Db 61 GCGGCTCGGCTGAGCGCGCGCAGCGCGCGCTCGCGGAGTGTGTCT 120
QY 842 GTTGTCTCTCTCGCTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 903
      |||
Db 121 GTTGTCTCTCTCGCTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 180
QY 942 CAGAGCGGACACTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 941
      |||
Db 121 CAGAGCGGACACTCTCTCGCGGAGGATACCGCGGAGTGTGTCT 218

RESULT 15
AL540382
LOCUS AL540382 LTI_FLO02_PL1 867 bp mRNA linear EST 16-FEB-2001
DEFINITION AL540382 LTI_FLO02_PL1 Homo sapiens cDNA clone CS0D001YK21 5 prime
, mRNA sequence.
```

```

ACCESSION AL540382
VERSION AL540382.1 GI:12870475
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
    source
        1..867
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="LII_FLO02_PL1"
            /lab_host="DH10B"
            /notes="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
            cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-stranded cDNA was digested with Not I
            and cloned into the Not I and Eco RV sites of the
            pCMVSPORT 6 vector. Library was constructed by Life
            Technologies. Contact : Feng Liang Life technologies, a
            division of Invitrogen 9800 Medical Center Drive Rockville
            , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
            fliang@lifetech.com URL :
            http://fulllength.invitrogen.com"
BASE COUNT 219 a 226 c 250 g 170 t 2 others
ORIGIN

Query Match 21.5%; Score 202; DB 9; Length 867;
Best Local Similarity 98.2%; Pred. No.3.9e-29;
Matches 224; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

Oy 716 TCGAGCGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGAGCG-6CGGCGGCTCGG 774
Db 1 TCGAGCGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGAGCGCGGCGCTCGG 60

Oy 775 AGCGGCTCCGGGCTCGGCTCGAGCGCCAGCGG-6CGGCTGGCGGAGGATTACCG 833
Db 61 AGCGGCTCCGGGCTCGGCTCGAGCGCCAGCGGCGGCTGGCGGAGGATTACCG 120

Oy 834 GCGAAGTGTTGTCTCTCGCTGGAGCCGCGAGAGCGGCGCTCAGGCGCGGCGCGCG 893
Db 121 GCGAAGTGTTGTCTCTCGCTGGAGCCGCGAGAGCGGCGCTCAGGCGCGGCGCGCG 180

Oy 894 GCGGCGAGGAGAGCGGACTCTGGCGCGGCTCTTTGGCGCGG 941
Db 181 GCGGCGAGGAGAGCGGACTCTKGCGCGGCTCTGTTGGCGCGG 228

Search completed: December 8, 2002, 12:49:39
Job time : 1237.21 secs

```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/930,548A
;; FILING DATE: 23-NOV-1992
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gordon, Alan M.
;; REGISTRATION NUMBER: 30,637
;; REFERENCE/DOCKET NUMBER: 31,298-01
;; TELEPHONE: 201-831-3244
;; TELEFAX: 201-831-3305
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4236 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..4068
;;

JS-07-930-548A-7

Query Match 25.3%; Score 76.2; DB 2; Length 4236;
Best Local Similarity 79.6%; Pred. No. 1.1e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 69 CTCGACTCTCTGGAAGTGATGCTCTGAAATCTATCTTTGACAAATCTACAGCACCA 128

DB 3194 CTCGCCCTCTTTGAAATGGAATGCCGCCAGAACAAATTTTGACAGACTGTACACAAATCC 3253

QY 129 AGACGACGTGGTGGTCTTACGAGTATGCTGTGGGAAATCTTCCTTAGGT 181

DB 3254 AGAGTGACGTGCTGTTTGGTCTTTTGTGTGGGAAATATTTTCCTTAGGT 3306

RESULT 5

US-07-813-593-3
; Sequence 3, Application US/07/813593
; Patent No. 5185438
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/813,593
; FILING DATE: 19920415
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601

;; REFERENCE/DOCKET NUMBER: LEM-3-PPP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-645-1405
;; TELEFAX: 212-645-2054
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5406 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 208..4311
;; NAME/KEY: Mat_peptide
;; LOCATION: 208..4308
;;

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTCTCTGGAATGGAATGCTCTGAAATCTATCTTTGACAAATCTACAGCACCA 128

DB 3395 CCGGACTCCCTTTGAGTGGATGCCGCCGGAACCAATTTTGACAGAGTATACACAAATTC 3454

QY 129 AGACGACGTGGTGGTCTTACGAGTATGCTGTGGGAAATCTTCCTTAGGT 181

DB 3455 AGACCAAGTGGTGGTCTTACGAGTATGCTGTGGGAAATATTTTCCTTAGGT 3507

RESULT 6

US-07-977-451-5
; Sequence 5, Application US/07/977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:

```

1 APPLICATION NUMBER: US 07/813,593
2
3 FILING DATE: 24-DEC-1991
4
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: US 07/793,065
7 FILING DATE: 15-NOV-1991
8
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: US 07/728,913
11 FILING DATE: 28-JUN-1991
12
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 07/679,666
15 FILING DATE: 02-APR-1991
16
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Feit, Irving N.
19 REGISTRATION NUMBER: 28,601
20 REFERENCE/DOCKET NUMBER: LEM-3-7P
21
22 TELECOMMUNICATION INFORMATION:
23 TELEPHONE: 212-645-1405
24
25 TELEFAX: 212-645-2054
26
27 INFORMATION FOR SEQ ID NO: 5:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 5406 base pairs
30 TYPE: NUCLEIC ACID
31 STRANDEDNESS: double
32 TOPOLOGY: linear
33
34 MOLECULE TYPE: cDNA
35
36 HYPOTHETICAL: NO
37
38 ANTI-SENSE: NO
39
40 FRAGMENT TYPE: N-terminal
41
42 FEATURE:
43 NAME/KEY: CDS
44 LOCATION: 208..4311
45
46 FEATURE:
47 NAME/KEY: mat_peptide
48 LOCATION: 265..4308
49
50 FEATURE:
51 NAME/KEY: sig_peptide
52 LOCATION: 208..264
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
```

[illegible]

RESULT 7
US-07-946-507-3
; Sequence 3, Application US/07946507
; Patent No. 5283354
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0. Version #1.25

```

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/946,507
  FILING DATE: 19920917
  CLASSIFICATION: 536
PRIOR APPLICATION DATA:
  PRIOR APPLICATION NUMBER: US/07/813,593
  FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/07/793,065
  FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/07/728,913
  FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/07/675,666
  FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Feil, Irving N.
  REGISTRATION NUMBER: 28,601
  REFERENCE/DOCKET NUMBER: LEM-3-PPP
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 212-645-1405
  TELEFAX: 212-645-2034
  INFORMATION FOR SEQ ID NO: 3:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 5406 base pairs
      TYPE: NUCLEIC ACID
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 208..4311
    FEATURE:
      NAME/KEY: mat_peptide
      LOCATION: 208..4308
  SS-07-946-507-3

```

```

Query Match      24.8%  Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%  Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY  59 CTCGACTTCCTGGAATGGATGGCTCCTGAATCTATCTTTGACAAATCTACGACCA 128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  3395 CCCGACTCCCTTTGAAGTGGATGGCCCCGGAACCACTTTTGACAGAGTATACACAAATC 3454

QY  129 AGACGCAGCTGGTGCTTACCGAGTATTCCTGTGGGAATCTTCCTTAGT 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DE  3455 AGACGCAAGTGGTCTTCGGTGTTGCTCTGGGAATATTCCTTAGT 3507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

1  RESULT 8
2  US-08-252-517-5
3  : Sequence 5, Application US/08252517
4  : Patent No. 5518065
5  :
6  : GENERAL INFORMATION:
7  :
8  : APPLICANT: Lemischka, Ihor R.
9  : TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
10 :
11 : TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
12 :
13 : NUMBER OF SEQUENCES: 10
14 :
15 : CORRESPONDENCE ADDRESS:
16 :
17 : ADDRESSEE: InClone Systems Incorporated
18 :
19 : STREET: 180 Varick Street
20 :
21 : CITY: New York
22 :
23 : STATE: New York
24 :
25 : COUNTRY: U.S.A.
26 :
27 : ZIP: 10014
28 :
29 : COMPUTER READABLE FORM:
30 :
31 : MEDIUM TYPE: Floppy disk
32 :
33 : COMPUTER: IBM PC compatible
34 :
35 : OPERATING SYSTEM: PC-DOS/MS-DOS
36 :
37 : SOFTWARE: Patent In Release #1.25
38 :
39 : CURRENT APPLICATION DATA:

```


APPLICATION NUMBER: US/08/252,517
FILING DATE: 31-OCT-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 19-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
US-08-252-517-5

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTCGACTTCCTCGAATGGATGGCTCTGAATCTATCTTTGACAAATCTACAGCACCA 128
DB 3395 CCCGACTCCCTTGAAGTGGATGGCCCGGAACCATTTTGACAGATACACATTC 3454
QY 129 AGAGCGAGTGTGCTTACGAGATTTGCTGGGAAATCTTCTCTTAGGT 181
DB 3455 AGAGCGAGTGTGCTTCTCGGTGTGCTGCTCGGTGGAAATATTTCCITAGGT 3507
RESULT 9
US-07-906-397A-5

Sequence 5, Application US/07906397A
Patent No. 5621090
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,397A
FILING DATE: 19920626
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 208..4308
US-07-906-397A-5

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTCGACTTCCTCGAATGGATGGCTCTGAATCTATCTTTGACAAATCTACAGCACCA 128
DB 3395 CCCGACTCCCTTGAAGTGGATGGCCCGGAACCATTTTGACAGATACACATTC 3454
QY 129 AGAGCGAGTGTGCTTACGAGATTTGCTGGGAAATCTTCTCTTAGGT 181
DB 3455 AGAGCGAGTGTGCTTCTCGGTGTGCTGCTCGGTGGAAATATTTCCITAGGT 3507

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 04:25:08 ; Search time 330.95 Seconds
(without alignments)
2048.198 Million cell updates/sec

Title: US-09-778-900A-3_COPY_200_500
Perfect score: 301
Sequence: 1 ggaactcgggttgatgtg.....taacaaaaaacctcttaagt 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
tal number of hits satisfying chosen parameters: 4370473

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :	N_Geneseq_101002.*
1:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	301	100.0	726	22	AAD19454 Human flt-1 gene D
2	114.4	38.0	4017	24	ABL91752 Human polynucleoti
3	112.8	37.5	7680	23	AAS70286 DNA encoding novel
4	76.2	25.3	4044	24	ABL91751 Human polynucleoti
5	76.2	25.3	4071	13	AAQ28272 A novel type III R
6	76.2	25.3	4071	20	AAV99829 Human receptor tyr
7	76.2	25.3	4071	24	ABL91754 Human polynucleoti
8	76.2	25.3	4225	22	RAF83308 Human VEGFR-2 enco
9	76.2	25.3	4236	19	AAV34763 Human KDR genomic

10	74.6	24.8	5404	17	AAT38735 Murine foetal live
11	74.6	24.8	5404	20	AAV77516 Murine flk-1 cDNA.
12	74.6	24.8	5406	13	AAQ29957 flk-1 cDNA sequenc
13	74.6	24.8	5406	14	AAQ53504 Murine flk-1 cDNA.
14	74.6	24.8	5406	14	AAQ35251 Human flk-1 coding
15	74.6	24.8	5406	14	AAQ40316 Murine flk-1 cDNA.
16	74.6	24.8	5406	16	AAQ81014 Flk1 receptor prot
17	74.6	24.8	5406	16	AAQ79070 Mouse flk-1 cDNA.
18	74.6	24.8	5406	18	AAI72119 Murine flk-1 recep
19	74.6	24.8	5470	15	AAQ64049 Sequence of murine
20	67.2	22.3	2949	19	AAV39041 Human receptor typ
21	67.2	22.3	2958	19	AAV39042 Human receptor typ
22	67.2	22.3	2958	19	AAV39038 Human receptor typ
23	67.2	22.3	2982	19	AAV39039 Human receptor typ
24	67.2	22.3	3120	14	AAQ49756 pTK gene lptK25.
25	67.2	22.3	3120	16	AAQ703096 Protein tyrosine-k
26	67.2	22.3	3475	16	AAI00802 Human FLK2/flt3 ty
27	67.2	22.3	3476	16	AAQ91536 Human STK-1 cDNA.
28	67.2	22.3	3501	14	AAQ53503 Human flk-2 cDNA.
29	67.2	22.3	3501	14	AAQ35250 Human flk-2 coding
30	67.2	22.3	3501	14	AAQ35250 Flk2 receptor prot
31	67.2	22.3	3501	16	AAQ81013 Murine flk-2 cDNA.
32	67.2	22.3	3501	16	AAQ79069 Murine foetal live
33	67.2	22.3	3501	17	AAI38734 Human flk-2 recept
34	67.2	22.3	3501	18	AAI72117 Human flk-2 cDNA.
35	67.2	22.3	3501	20	AAI77515 Human polynucleoti
36	67.2	22.3	3501	24	ABL91753 Human Flt4/VEGFR-3
37	65.6	21.8	3897	21	AAQ62405 Human Flt4/VEGFR-3
38	65.6	21.8	4111	22	AAQ68952 FLT4 receptor tyro
39	65.6	21.8	4111	17	AAT12068 Human Flt4 recepto
40	65.6	21.8	4195	21	AAA37815 Human tyrosine kin
41	65.6	21.8	4195	21	AAQ52333 Protein tyrosine-k
42	65.6	21.8	4425	16	AAQ03090 Nucleotide sequenc
43	65.6	21.8	4450	21	AAQ62210 Human Flt4 recepto
44	65.6	21.8	4795	21	AAA37816
45	65.6	21.8	4795	21	AAA37816

ALIGNMENTS

RESULT 1	
AAI15454	
ID	AAI15454 standard; DNA; 726 BP.
XX	AC AAD19454;
XX	AC AAD19454;
DT	18-DEC-2001 (first entry)
XX	Human flt-1 gene DNA fragment #3.
DE	Human flt-1 gene; cytostatic; haemostatic; gene mapping; gynaecological;
KW	optalmological; rheumatoid arthritis; endometriosis; angiogenic disease;
KW	diabetic retinopathy; psoriasis; VEGFR-1; drug therapy; pharmacogenetic;
KW	vascular endothelial growth factor; single nucleotide polymorphism; SNP;
KW	cancer; medication; ds.
XX	Homc sapiens.
XX	Homc sapiens.
XX	Homc sapiens.
Key	Location/Qualifiers
FT intron	1..286
FT exon	/tag= a
FT intron	267..278
FT intron	/tag= b
FT intron	279..726
FT intron	/tag= c
FT variation	replace (454, A)
FT variation	/tag= d
FT variation	/standard_name= *Single nucleotide polymorphism (SNP)*
EP11.0123-A2.	
XX	05-SEP-2001.
PD	05-SEP-2001.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW Homo sapiens.
OS W0200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG06099.
XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity
XX Claim 1; SEQ ID No 6090; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Query Match 37.5%; Score 112.8; DB 23; Length 7680;
Best Local Similarity 98.3%; Pred. No. 1.1e-25;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 66 AGACTGACCTCTCTGAAATGGATGGCTCTCTGAAATCTATCTTTGACAAATCTCTACGCA 125
DB 3422 ATACTGACCTCTCTGAAATGGATGGCTCTCTGAAATCTATCTTTGACAAATCTCTACGCA 3481
QY 126 CCAAGACGACGTGGCTTACGAGTATGCTGTGGGAATCTCTCTTAGGT 181
DB 3482 CCAAGACGACGTGGCTTACGAGTATGCTGTGGGAATCTCTCTTAGGT 3537
RESULT 4
ABL91751
ID ABL91751 standard; DNA; 4044 BP.
XX ABL91751;
XX 28-MAY-2002 (first entry)
DT

XX Human polynucleotide SEQ ID NO 94.
DE Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen; Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide; cytostatic; virucide; protozoacide; antibacterial; ds.
XX Homo sapiens.
XX DE10100586-C1.
XX 11-APR-2002.
XX 09-JAN-2001; 2001DE-1000586.
XX 09-JAN-2001; 2001DE-1000586.
XX (RIBO-) RIBOPHARMA A3.
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX WPI; 2002-270454/32.
XX Inhibiting gene expression in cells, useful for e.g. treating tumors, by introducing double-stranded complementary oligoRNA having unpaired terminal bases
XX Claim 13; Page 67-68; 104pp; German.
XX The invention relates to a method for inhibiting expression of a target gene (ABL51658-ABL51797) in a cell by introducing at least one oligoribonucleotide that has a double-stranded structure consisting of at most 49 sequential nucleotide pairs, with at least part of one strand complementary with the target gene and has at least one end a single-stranded segment of 1-4 nt. The method provides oligoribonucleotides for antisense inhibition of gene expression useful e.g. for treating tumors but the oligoribonucleotides may also be directed against genes present in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans, animals or plants) or against cytokine, Id, developmental or prion genes. The method provides more effective inhibition of gene expression than use of known oligonucleotides, probably because the unpaired overhang increases stability and thus intracellular concentration.
XX Query Match 25.3%; Score 76.2; DB 24; Length 4044;
Best Local Similarity 79.6%; Pred. No. 5e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 69 CTCGACTTCTCTGAAATGGATGGCTCTCTGAAATCTATCTTTGACAAATCTCTACGCA 128
DB 3194 C1G3CCTCCCTTGAATGGATGGCTCTCTGAAATCTATCTTTGACAAATCTCTACGCA 3253
QY 129 AGACGACGCTGGCTCTTACGAGTATGCTGTGGGAATCTCTCTTAGGT 181
DB 3254 AGAGTGACGCTGGCTCTTGTGGTTTGTGGGAATCTCTCTTAGGT 3306
RESULT 5
AAQ28272
ID AAQ28272 standard; cDNA; 4071 BP.
XX AAQ28272;
XX 15-FEB-1993 (first entry)
XX A novel type III RIK gene - the KDR gene.
XX Receptor tyrosine kinase; vascular endothelial cell growth factors; cancer; tumour; diagnosing; monitoring; ss.
XX Homo sapiens.
OS

```
XX FH Key Location/Qualifiers
XX CDS 1..4068
XX FT /*tag= a
XX PN W09214748-A.
XX PD 03-SEP-1992.
XX PF 20-FEB-1992; 92WO-US01300.
XX PR 22-FEB-1991; 91US-0657236.
XX PS (AMCY ) AMERICAN CYANAMID CO.
XX PI Carrión ME, Terman BI;
XX DR NPI; 1992-316117/38.
XX DR P-PSDB; AAR26999.
XX PS DNA encoding type III receptor tyrosine kinase - useful for
    diagnosing the onset of cancer
XX Claim 3; Fig 7; 101pp; English.
XX This sequence represents a novel type III receptor tyrosine kinase
XX gene. A labelled EcoRI-BamHI DNA segment derived from clone Bf111081.8
XX was used as a probe to rescreen a human endothelial cDNA library
XX (HLI0246) for 5' full length DNA segments of the gene from which the
XX insert portion of Bf111081.8 is derived. A synthetic probe designed
XX from nucleotides 3297-3325 of Bf111081.8 is then used to isolate
XX more 3' full length clones. One of the clones, designated Bf111200.2
XX is cloned into pBluescript KS and the synthetic oligonucleotide
XX JGACGCCG ATC GAG cloned, which contains the initial sequence Met-
XX Glu, the first two amino acids encoded by the KDR gene, forming
XX Bf1140, which is then purified on a CsCl density gradient. This
XX was sequenced, together with Bf111081.8, and Bf11129.5 to comprise
XX the entire ORF of 4,068 nucleotides of the KDR gene.
XX Sequence 4071 BP; 1172 A; 894 C; 1024 G; 981 T; 0 other;
XX Query Match 25.3%; Score 76.2; DB 13; Length 4071;
XX Best Local Similarity 79.6%; Pred. No. 5e-14;
XX Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
XX QY 69 CTCGACTTCTCTGAAATGGATGGCTCTGAAATCTATCTTTCACAAATCTACAGCACCA 128
XX DB 3154 CTCGCTCCCTTGAATGGATGGCTCTGAAATCTATCTTTCACAAATCTACAGCACCA 3253
XX 129 AGAGCGACGTGTGGTCTTACGAGATATGCTGTGGGAAATCTCTCTTAGGT 181
XX DB 3254 AGAGTACGCTGTGGTCTTACGAGATATGCTGTGGGAAATATTTCTTAGGT 3306
XX RESULT 6
XX ID AAV99829 standard; cDNA; 4071 BP.
XX AC AAV99829;
XX DT 12-APR-1999 (first entry)
XX DE Human receptor tyrosine kinase KDR cDNA.
XX KDR; receptor tyrosine kinase; human; signal transduction;
XX mitogen; neoplasms; angiogenesis; diabetes; retinopathy;
XX breast cancer; brain cancer; inflammation; rheumatoid arthritis;
XX psoriasis; contact dermatitis; hypersensitivity; antagonist;
XX gene therapy; ds.
XX Homo sapiens.
XX W09858053-A1.
```

```
XX PD 23-DEC-1998.
XX PF 17-JUN-1998; 98WO-US12569.
XX PR 18-JUN-1997; 97US-0050962.
XX PA (MERI ) MERCK & CO INC.
XX PI Kendall RL, Mao X, Tebben A, Thomas KA;
XX DR NPI; 1999-095333/08.
XX DR P-PSDB; AAW80997.
XX PS Human receptor tyrosine kinase protein, KDR - useful e.g. to screen
    for antagonists useful to treat diseases involving neovascularization
    e.g. diabetic retinal vascularization, cancers
XX Claim 1; Fig 1a; 69pp; English.
XX This nucleotide sequence encodes a novel receptor tyrosine kinase,
XX termed KDR (see AAW80997), that is expressed on human endothelial
XX cells. KDR is activated by vascular endothelial growth factor and
XX mediates a mitogenic signal. It is implicated in clinical
XX neovascularization. KDR cDNA was isolated from a human umbilical vein
XX endothelial cell lambda phage cDNA library using a 576-bp DNA probe
XX that had been generated by PCR (see also AAW9850-53). The predicted
XX protein product has amino acid differences from the previously
XX published KDR sequence at positions 498 (Ala to Glu), 772 (Thr to
XX Ala), 787 (Gly to Arg), 835 (Asn to Lys), 848 (Glu to Val) and 1347
XX (Thr to Ser), producing a protein predicted by computer modeling to
XX have higher activity and functionality. The invention also relates
XX to recombinant vectors and recombinant hosts which contain a DNA
XX fragment encoding human KDR, a DNA fragment encoding the
XX intracellular portion of KDR with or without a membrane anchor
XX sequence, purified forms of associated human KDR, and human mutant
XX forms of KDR. KDR, fusion proteins or fragments can be used in
XX assays to identify antagonists and agonists of human KDR (claimed).
XX Antagonists of KDR useful for treating diseases involving
XX neovascularization e.g. diabetic retinal vascularization, cancers
XX (e.g. brain, breast, etc.) and forms of inflammation e.g. rheumatoid
XX arthritis, psoriasis, contact dermatitis and hypersensitivity
XX reactions. The polynucleotides are useful to screen for KDR
XX antagonists/agonists and for gene therapy (e.g. by introducing a
XX gene portion encoding a KDR protein containing functional ligand
XX binding and membrane anchoring moieties but not tyrosine kinase
XX activity). They are also useful to measure levels of human KDR.
XX Sequence 4071 BP; 1169 A; 894 C; 1027 G; 981 T; 0 other;
XX Query Match 25.3%; Score 76.2; DB 20; Length 4071;
XX Best Local Similarity 79.6%; Pred. No. 5e-14;
XX Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
XX QY 69 CTCGACTTCTCTGAAATGGATGGCTCTGAAATCTATCTTTCACAAATCTACAGCACCA 128
XX DB 3154 CTCGCTCCCTTGAATGGATGGCTCTGAAATCTATCTTTCACAAATCTACAGCACCA 3253
XX QY 129 AGAGCGACGTGTGGTCTTACGAGATATGCTGTGGGAAATCTCTCTTAGGT 181
XX DB 3254 AGAGTACGCTGTGGTCTTACGAGATATGCTGTGGGAAATATTTCTTAGGT 3306
XX RESULT 7
XX ID ABL91754 standard; DNA; 4071 BP.
XX AC ABL91754;
XX DT 28-MAY-2002 (first entry)
XX DE Human polynucleotide SEQ ID NO 97.
```



```
XX PN US5766860-A.
XX PD 16-JUN-1998.
XX PF 25-FEB-1997; 97US-0810116.
XX PR 23-NOV-1997; 92US-0930548.
XX PR 25-FEB-1997; 97US-0810116.
XX PA (AMCY ) AMERICAN CYANAMID CO.
XX PI Carrion ME, Terman BI;
XX DR WPI; 1998-361682/31.
XX DR P-PSDB; AAW59275.
XX PT Screening assay for vascular endothelial cell growth factor
XX PT antagonists - using recombinant cells expressing receptor protein
XX PT Claim 1; Fig 7A-M; 51pp; English.
XX - This sequence encodes a novel human growth factor receptor, kinase
CC insert domain containing receptor or KDR. This receptor is capable of
CC binding to the vascular endothelial cell growth factor, VEGF and is
CC used in a screening assay which identifies compounds that inhibit VEGF
CC action on KDR. Such compounds which inhibit binding of VEGF to the KDR
CC may inhibit angiogenesis and thus be useful for treating cancer.
XX SQ Sequence 4236 BP; 1216 A; 938 C; 1062 G; 1020 T; 0 other;
XX
Query Match 25.3%; Score 76.2; DB 19; Length 4236;
Best Local Similarity 79.6%; Pred No. 5.1e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 69 CTCGACTTCTCTGAATGGATGGCTCCCTGGAATCTATCTGGAATCTCTCTTAGGT 181
DB 3194 CTCGCTCCCTTGAATGGATGGCTCCCAAGAAATTTTGACAGAGTGACACAATCC 3253
QY 129 AGAGCAGCTGTGGTCTAGGAGTATTCTGTGGAAATCTCTCTTAGGT 181
DB 3254 AGAGTACCTGTGGTCTTTGGTGTGTTTGTGTGGAAATTTCTCTTAGGT 3306
RESULT 10
RAT38735
ID AAT38735 standard; cDNA; 5404 BP.
% C AAT38735;
11-DEC-1996 (first entry)
AX Murine foetal liver kinase 1 cDNA.
DE Murine foetal liver kinase 1; flk-1; protein tyrosine kinase;
KW monoclonal; antibody; extracellular domain; receptor assay;
KW haematopoietic stem cell; ligand; stimulation; proliferation;
KW differentiation; treatment; anaemia; bone marrow damage;
KW cancer chemotherapy; radiation; ds.
XX Mus musculus.
OS
XX Key Location/Qualifiers
XX CDS 208..4311
XX sig_peptide 208..264
XX mat_peptide 265..4308
XX FT /tag= a
XX FT /tag= b
XX FT /tag= c
XX PN US5548065-A.
XX PD 20-AUG-1996.
```

```
XX PF 02-APR-1991; 91US-0679666.
XX PR 15-NOV-1992; 92US-0977451.
XX PR 02-APR-1991; 91US-0679666.
XX PR 28-JUN-1991; 91US-0726913.
XX PR 15-NOV-1991; 91US-0793065.
XX PR 24-DEC-1991; 91US-0813593.
XX PR 26-JUN-1992; 92US-0906397.
XX PR 12-NOV-1992; 92US-0975049.
XX PR 30-APR-1993; 93US-0055269.
XX PR 31-OCT-1994; 94US-0252517.
XX (UYPR-) UNIV PRINCEJON.
XX PA Lemischka IR;
XX PI WPI; 1996-392678/39.
XX DR P-PSDB; AAT38735.
XX DR Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays,
XX PT for isolating haematopoietic stem cells expressing receptor and for
XX PT obtaining ligands
XX PS Disclosure; Columns 51-62; 50pp; English.
XX CC The present sequence encodes murine foetal liver kinase 1 (flk-1),
CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
CC raised against the extracellular portion of flk-1 can be used to
CC assay for flk receptors on the surface of haematopoietic stem
CC cells, and to isolate positive cells. The antibodies can also
CC be used as, or to obtain ligands, which stimulate the proliferation
CC and/or differentiation of stem cells. The ligands can be used, e.g.
CC for treating anaemia, or bone marrow damage resulting from cancer
CC chemotherapy, or radiation.
XX SQ Sequence 5404 BP; 1411 A; 1297 C; 1423 G; 1273 T; 0 other;
XX
Query Match 24.8%; Score 74.6; DB 17; Length 5404;
Best Local Similarity 78.8%; Pred. No. 1.8e-13;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTCGACTTCTCTGGAATGGATGGCTCCCTGGAATCTATCTGGAATCTCTCTTAGGT 181
DB 3395 CCGGACTCCCTTGAAGTGGATGGCTCCCGGAAACCATTTTGACAGAGTATACACAATC 3454
QY 129 AGAGCAGCTGTGGTCTACGAGTATTCTGTGGAAATCTCTCTTAGGT 181
DB 3455 AGAGCAGCTGTGGTCTTCTGCTGTGGTGTGTTCTGTGGAAATTTCTCTTAGGT 3507
RESULT 11
RAT38735
ID AAT38735 standard; cDNA; 5404 BP.
XX AC AAT38735;
XX DT 05-AUG-1999 (first entry)
XX DE Murine flk-1 cDNA.
XX KW Murine; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
XX KW monoclonal; polyclonal; antibody; tyrosine kinase; ds.
XX OS Mus sp.
XX PH Key Location/Qualifiers
XX CDS 208..4311
XX FT /tag= a
XX FT /product= "flk-1"
XX PN US5912133-A.
XX PD
```

Db 3687 CTCTCTGCAAGCAAAATCGCAGCAGGATGCAAGAACTATAT 3728

RESULT 15

AAQ79070
ID AAQ79070 standard: cDNA; 5406 BP.

XX AAQ79070;

XX 04-JUL-1995 (first entry)

XX Mouse flk-1 cDNA.

XX Fetal liver kinase-1; flk-1; protein tyrosine-kinase receptor;
XX hematopoiesis; stem cell; ds.

OS Mus sp.

Key Location/Qualifiers

CDS 208..4311

/*tag= a

sig_peptide 208..264

/*tag= b

mat_peptide 265..4308

/*tag= c

XX US5367057-A.

XX 22-NOV-1994.

XX 02-APR-1991; 91US-0679666.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 26-JUN-1992; 92US-0906397.

XX 12-NOV-1992; 92US-0975049.

XX 19-NOV-1992; 92US-0977451.

XX 30-APR-1993; 93US-0055269.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1995-005894/01.

XX P-PSDB; AAR67537.

XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate
XX proliferation and/or stimulation of primitive mammalian
XX hematopoietic stem cells in vitro or in vivo.

XX Disclosure: Fig. 3A-3I; 69pp; English.

XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
XX kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAR67535-70.
XX respectively, and the deduced amino acid sequences in AAR67535-37,
XX respectively.

XX Sequence 5406 BP; 1412 A; 1298 C; 1423 G; 1273 T; 0 other;

XX Query Match 33.7%; Score 708.4; DB 16; Length 5406;

XX Best Local Similarity 65.0%; Pred. No. 1.3e-204;

XX Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;

XX 201 CACTCTTAATCTTACCATCATGAATGTTTCCCTCGAAGATTCAGGCACCTATGCTCGAG 250

XX 2070 CACTTGATTTGGCTATTCAGAAATGCTCTCTCGAGGACCAAGGGGACTATGTTGTC 2129

XX 261 ACCGAGGATGATACACGGGGAAGAAATCTCCAGAGAGAGAAATACAAATCAGAGA 320

XX 2130 TGCTCAAGATAAGAGACCAAGAAAGACATGCTTGCTGCTCAACACGCTCACTCTTGA 2189

OY 321 TCAGGAAGCACCATACCTCTCGCAAACTCAGTCATCACACAGTGGCCATCAGCAGTTC 380
DB 2190 GCGCATGCCACCCATGATACCGGAATCTGGAGAACTCAGACAAACACCATGGCGAGAC 2249
OY 381 CACCACIIITAGACTGTCATGCTAATGTTGCCCGAGCCTCAGATCCTACATTTGTTAAAAA 440
DB 2250 CATTGAAGTACTTGGCCAGCATCTGGAATCCTACCCACACATTAATGTTCAAGA 2309
OY 441 CAACCAAAATACAAAGAGCCTGGAATTTATTTAGGACCCAGAGCAGCAGCTGTT 500
DB 2310 CAAGGAGACCTGGTAGAAGATTTCAGGCATTTGCTAGAGATGGGAACCGAACCTGAC 2369
OY 501 TATTGAAGTACAGAGAGGATGAAGTGTCTATCACTGCAAGCCACCAACAGAA 560
DB 2370 TATCCCGAGGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2429
OY 561 GGGCTGTGGAAAGTTCAGCATACCTCAGTGTTCAGGAACCTCGGACAAAGTCTAATCT 620
DB 2430 TGGCTGTGCAAGCGGAGAGCGCTCTTCATATAGAAGTGCCTCCAGGAAAGACCACT 2489
OY 621 GGAGCTGATCAGCTTACATGACCTGTGGCTGGGACTCTCTTCTGGCTCTCTATTAAC 580
DB 2490 GGAAGTCAITATCTCTGTCGGCAGTGCAGTGAATGCCATGTTCTCTGGCTCTCTTGT 2549
OY 681 CTGCTTATCCGAAATGAAAAGTCTCTTCTGAAATAAGAGTCTGACTACCTATC 737
DB 2550 CATTCTGTACGACCGGTTAGCGGGCCATGAAGGGAACTGAGAGCAGGCTACTGTCT 2609
OY 738 AATTATATGACCCAGATGAAGTTCCTTTGGATGAGAGTGTGAGCGGCTCCCTTATGA 797
DB 2610 TATTGTCATGATCCAGATGAATGCGCTTGGATGAGCGCTGTGAACCTTGGCTTATGA 2669
OY 798 TGCAGCAGAGTGGAGTTCGCGGAGAGACTTAACTGGGCAATCAGTGTGAAGAG 857
DB 2670 TGCAGCAGAGTGGAGTTCGCGGAGAGACTTAACTGGGCAATCAGTGTGAAGAG 2729
OY 858 GGCCTTTGGAAGAGTGTTCAGCATCAGCAATTTGGCAATTAAGAAATCAGTGTGCG 917
DB 2730 TGCCTTCGCGCAGTGAATGAGGAGAGCGCTTTTGAATGCAAGAGCAGGCTTGA 2789
OY 918 GACTGTGGCTGTGAAGTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
DB 2790 AACAGTAGCGCTCAAGATGTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2849
OY 978 GACTGAGCTAAATCTTGACCCACATTTGCCACCATCTGAAGCTGTGTTAACTGTGG 1037
DB 2850 GTCTGAACTCAAGTCTTCCATCCATTTGGTCACTATCTATGTTGAGGAGGAGGAG 2909
OY 1038 AGCCTGCACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097
DB 2910 GCGCTGCACCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2969
OY 1098 TCTCTCACTACTCAAGAGCAAGCTGACTTATTTTCTCAACAGGATGAGCAGCT 1157
DB 2970 CCTATCACTACTTACGGGCAAGAGAAATGAATTTGTTCCCTATAGAGCAAGAGGAG 3029
OY 1158 ACACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1217
DB 3030 AGGCTTCGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3086
OY 1218 ACTAGTAGGCTCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1277
DB 3087 CTTCAGAGCAATCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3146
OY 1278 TGTGAGTGTGTTGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1337
DB 3147 GCTCAGTGTGAGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3206
OY 1338 GGAGATCTGATTTTACAGTTTTCAAGTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1397
DB 3207 GGAGCATCTCATGTTTACAGTTTCCAAAGTGGCTAAGGAGGAGGAGGAGGAGGAGGAG 3266
OY 1398 AAGTGTCAITCATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1457

Db 3267 GAAGTGTATCCACAGGACCTGGCAGCAGAAACATTCCTATCGGAGAAATGTSGT 3326
QY 1458 GAAGATTGTGATTTTGGCCCTTCCCGGATATTTATACAAACCCGATTAIGTCAGAA 1517
Db 3327 TAAGATCTGTGACTTGGCTTGGCCCGGACATTTATNAAGACCCGATTAIGTCAGAA 3386
QY 1518 AGGAGATACTCGACTTCCTCTGAAATGATGGCTCCGAACTATCTTTTGACAAATCTA 1577
Db 3387 AGGAGATCCCGACTCCCTTTGAAGTGAATGGCCCGGAACCATTTTGCACAGTATA 3446
QY 1578 CAGCACACAGAGAGAGCTGTGCTTACGGAGTATTCCTGTGGAAATCTTCTCTTAGG 1637
Db 3447 CACAATTCAGAGAGATGTGTGCTTTCGGGTGTTGCTCTGGAAATATTTCTTAGG 3506
QY 1638 TGGGTCTCCATACCCAGGAGTACAAATGGATGAGGACTTTTGCAGTCCCTGAGGGAAGG 1697
Db 3507 TCCCTCCCATCCCTGGGTCAGATTAAGAGATTTTGTAGAGATTAAGAGAGG 3566
QY 1698 CATGAGATGAGAGCTCCTGTAGTACTTACTCCTGAAATCTATCAGATCATCTGGACTG 1757
Db 3567 AACTAGATGGGGCTCCTGACTACACTACCCAGAAATGTACCAGACCATCTGGACTG 3626
QY 1758 CTGGCACAGACCCCAAGAGAGCCCAAGATTTGAGAGACTTTGCGAAAACTAGGTGA 1817
Db 3627 CTGGCATGAGACCCCAACAGAGACCTCGTTTTACAGATTGGTGGAGCTATTGGGAAA 3686
QY 1818 TTGCTTCAAGCAAAATGTACACAGAGATCGTAAAGACTACAT 1859
Db 3687 CCTCTGCAAGCAAAATGCGCAGCAGAGATGGCAAGAGACTATAT 3728

Search completed: December 8, 2002, 09:38:18
Job time : 518.686 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 8, 2002, 07:52:47 ; Search time 2363.64 Seconds
(without alignments)
11586.276 Million cell updates/sec

Title: D64016_COPY_500_1440
Perfect score: 941
Sequence: 1 GCCCAAAACACACGACAC.....CCCCGGTCTTTGGCCGGCGG 941

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pi.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vt.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vt.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_higo_hum.*
- 40: em_higo_mus.*
- 41: em_higo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	941	100.0	1745	6	AR157385	AR157385 Sequence
2	941	100.0	1745	9	HUMES	D64016 Human gene
3	939.4	99.8	102119	9	ALI39005	ALI39005 Human DNA
4	667	70.9	840	6	E10168	E10168 Promoter of
5	613	65.1	6316	6	AX251105	AX251105 Sequence
6	613	65.1	6316	6	AX351868	AX351868 Sequence
7	613	65.1	6316	6	AX344256	AX344256 Sequence
8	613	65.1	6316	6	AX348647	AX348647 Sequence
9	613	65.1	6316	6	AX458547	AX458547 Sequence
10	510.8	54.3	6316	6	AX251106	AX251106 Sequence
11	510.8	54.3	6316	6	AX251869	AX251869 Sequence
12	510.8	54.3	6316	6	AX344257	AX344257 Sequence
13	510.8	54.3	6316	6	AX348648	AX348648 Sequence
14	510.8	54.3	6321	6	AX458548	AX458548 Sequence
15	375.4	39.9	174496	2	AC119074	AC119074 Rattus no
16	361.2	38.4	232575	2	AC122299	AC122299 Mus muscu
17	359.6	38.2	2648	10	MMU224863	AJ224863 Mus muscu
18	194.8	20.7	2313	6	BD003577	BD003577 Gene ther
19	194.8	20.7	2313	6	I82801	I82801 Sequence 5
20	194.8	20.7	2523	6	AR163540	AR163540 Sequence
21	194.8	20.7	2523	6	AR163540	E13256 Human mRNA
22	194.8	20.7	2523	6	E13256	E14000 Human mRNA
23	194.8	20.7	2523	6	E14000	AR030842 Sequence
24	194.8	20.7	2651	6	AR030842	U01134 Human solub
25	194.8	20.7	2651	9	U01134	X51602 Human flt m
26	194.8	20.7	7680	6	HSFLI	AX060540 Sequence
27	130.6	13.9	3009	6	AX060540	AC130633 Rattus no
28	78.2	8.3	159184	2	AC130633	AC129765 Rattus no
29	76.6	8.1	185263	2	AC129765	AC118303 Rattus no
30	76.2	8.0	174031	2	AC118303	AC125560 Rattus no
31	75.6	7.9	95282	2	AC125560	AC126528 Rattus no
32	74.5	7.9	95282	2	AC126528	AC103544 Rattus no
33	74.2	7.9	122113	2	AC103544	AC127844 Rattus no
34	73	7.8	292136	2	AC127844	AC099104 Rattus no
35	72.4	7.7	125159	2	AC099104	AC129739 Rattus no
36	72	7.7	174303	2	AC129739	AC121738 Rattus no
37	71	7.5	27913	2	AC121738	AC118303 Rattus no
38	70.8	7.5	111079	2	AC118303	AC114443 Rattus no
39	70.8	7.5	174031	2	CNS08CA9	AL831796 Oryza sat
40	70.6	7.5	150238	2	AC105514	AC087563 Homo sapi
41	70.6	7.5	155406	2	AC087563	AC121481 Rattus no
42	70.6	7.5	298166	2	AC121481	
43	70.4	7.4	85022	2	AC121481	
44	70	7.4	123071	2	AC121481	
45	70	7.4	123071	2	AC121481	

ALIGNMENTS

RESULT 1
AR157385
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AR157385
Sequence 1 from patent US 6245512.
AR157385
AR157385.1 GI:16218320
Unknown.
Unclassified.
1 (bases 1 to 1745)
Williams, L.I. and Morishita, K.
Promoter for VEGF receptor
Patent: US 6245512-A 1 12-JUN-2001;
Location/Qualifiers

1745 bp
DNA
linear
PAT 17-OCT-2001

Filed 3/15/99
date of filed
11/9/95

Pred. No. is the number of results predicted by chance to have a

```
source 1. 1745
BASE COUNT 315 a 501 c 621 g 308 t
ORIGIN
Query Match 100.0%; Score 941; DB 6; Length 1745;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 941; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCCGAAACACACGACGACGCTCCCTGGACCTGAGTGGTTCGACGCTTCCCAAG 60
DB 500 GCCGAAACACACGACGACGCTCCCTGGACCTGAGTGGTTCGACGCTTCCCAAG 559
QY 61 GTGCCAAGACGCTCAGTTCCTCCAGGCTCCAGGCTCCAGGCTTGTGCGGAGGT 120
DB 560 GTGCCAAGACGCTCAGTTCCTCCAGGCTCCAGGCTTGTGCGGAGGT 619
QY 121 CTCGGTGCCTTCTAGACTTCGCGGACGCTGAGGGGTCGAGGGGTCAGGCGACG 180
DB 620 CTCGGTGCCTTCTAGACTTCGCGGACGCTGAGGGGTCGAGGGGTCAGGCGACG 679
QY 181 GCGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 240
DB 680 GCGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 739
QY 241 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 740 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 799
QY 301 GCGGAGCTTACGCTTGTCCCTTCCAGTTCGCGGCGGCGGCGGCGGCGGCGGCG 360
DB 800 GCGGAGCTTACGCTTGTCCCTTCCAGTTCGCGGCGGCGGCGGCGGCGGCGGCG 859
QY 361 GGTGAGGAGGCTGCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 420
DB 860 GGTGAGGAGGCTGCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 919
QY 421 AGAGGCGCGGAGCAAGACGCTGACCTGCGGGGCGGCGGCGGCGGCGGCGGCGG 480
DB 920 AGAGGCGCGGAGCAAGACGCTGACCTGCGGGGCGGCGGCGGCGGCGGCGGCGG 979
QY 481 AGACCTTCCGACGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 540
DB 980 AGACCTTCCGACGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1039
QY 541 CCCTAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 1040 CCCTAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
QY 601 GGTGCGGAGGCTTCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 660
DB 1100 GGTGCGGAGGCTTCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 1159
QY 661 AGGTATATAATCGCCGCGGCTGCGGCTGCTTATCATGAGGTCGCGGCGGCGGCGG 720
DB 1160 AGGTATATAATCGCCGCGGCTGCGGCTGCTTATCATGAGGTCGCGGCGGCGGCGG 1219
QY 721 GCGCGCGGCGGAGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCT 780
DB 1220 GCGCGCGGCGGAGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCT 1279
QY 781 CTCGGGGCTGGGTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 1280 CTCGGGGCTGGGTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1339
QY 841 GGTGTCTCTGCTGAGGCGGCGGAGACGCGGCGGCTCAGGCGGCGGCGGCGGCGG 900
DB 1340 GGTGTCTCTGCTGAGGCGGCGGAGACGCGGCGGCTCAGGCGGCGGCGGCGGCGG 1399
QY 901 ACAGAGGAGGAGCTTCTGCGGCGGCGGCTTTTGGCGGCGG 941
DB 1400 ACAGAGGAGGAGCTTCTGCGGCGGCGGCTTTTGGCGGCGG 1440
```

```
RESULT 2
HOMES 1745 bp DNA linear PRI 14-APR-2000
LOCUS Human gene for vascular endothelial growth factor receptor,
DEFINITION Human gene for vascular endothelial growth factor receptor,
ACCESSION D64016.1 GI:1088437
VERSION D64016.1
KEYWORDS vascular endothelial growth factor receptor,
SOURCE Homo sapiens placenta DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Morishita, K., Johnson, D.E. and Williams, L.T.
A novel promoter for vascular endothelial growth factor receptor
(flt-1) that confers endothelial-specific gene expression
J. Biol. Chem. 270 (46), 27948-27953 (1995)
JOURNAL 96070934
MEDLINE 2000(bases 1 to 1745)
REFERENCE 1 (bases 1 to 1745)
AUTHORS Morishita, K., Johnson, D.E. and Williams, L.T.
TITLE A novel promoter for vascular endothelial growth factor receptor
(flt-1) that confers endothelial-specific gene expression
JOURNAL J. Biol. Chem. 270 (46), 27948-27953 (1995)
FEATURES
source
location/Qualifiers
1. 1745
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="Placenta"
/evidence=not_experimental
1154..1171
exon
1196..1425
/product="vascular endothelial growth factor receptor"
/notes="flt-1"
/number=1
/evidence=experimental
1426..1745
intron
BASE COUNT 315 a 501 c 621 g 308 t
ORIGIN
Query Match 100.0%; Score 941; DB 9; Length 1745;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 941; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGAAACACACGACGACGCTCCCTGGACCTGAGTGGTTCGACGCTTCCCAAG 60
DB 500 GCCGAAACACACGACGACGCTCCCTGGACCTGAGTGGTTCGACGCTTCCCAAG 559
QY 61 GTGCCAAGACGCTCAGTTCCTCCAGGCTCCAGGCTCCAGGCTTGTGCGGAGGT 120
DB 560 GTGCCAAGACGCTCAGTTCCTCCAGGCTCCAGGCTTGTGCGGAGGT 619
QY 121 CTCGGTGCCTTCTAGACTTCGCGGACGCTGAGGGGTCGAGGGGTCAGGCGACG 180
DB 620 CTCGGTGCCTTCTAGACTTCGCGGACGCTGAGGGGTCGAGGGGTCAGGCGACG 679
QY 181 GCGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 240
DB 680 GCGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 739
QY 241 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 740 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 799
QY 301 GCGGAGCTTACGCTTGTCCCTTCCAGTTCGCGGCGGCGGCGGCGGCGGCGGCG 360
DB 800 GCGGAGCTTACGCTTGTCCCTTCCAGTTCGCGGCGGCGGCGGCGGCGGCGGCG 859
QY 361 GGTGAGGAGGCTGCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 420
DB 860 GGTGAGGAGGCTGCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 919
QY 421 AGAGGCGCGGAGCAAGACGCTGACCTGCGGGGCGGCGGCGGCGGCGGCGGCGG 480
DB 920 AGAGGCGCGGAGCAAGACGCTGACCTGCGGGGCGGCGGCGGCGGCGGCGGCGG 979
QY 481 AGACCTTCCGACGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 540
DB 980 AGACCTTCCGACGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1039
QY 541 CCCTAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 1040 CCCTAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
QY 601 GGTGCGGAGGCTTCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 660
DB 1100 GGTGCGGAGGCTTCAAGGATTTCTGAGCGGATGCGGCGGCGGCGGCGGCGGCGG 1159
QY 661 AGGTATATAATCGCCGCGGCTGCGGCTGCTTATCATGAGGTCGCGGCGGCGGCGG 720
DB 1160 AGGTATATAATCGCCGCGGCTGCGGCTGCTTATCATGAGGTCGCGGCGGCGGCGG 1219
QY 721 GCGCGCGGCGGAGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCT 780
DB 1220 GCGCGCGGCGGAGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCT 1279
QY 781 CTCGGGGCTGGGTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 1280 CTCGGGGCTGGGTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1339
QY 841 GGTGTCTCTGCTGAGGCGGCGGAGACGCGGCGGCTCAGGCGGCGGCGGCGGCGG 900
DB 1340 GGTGTCTCTGCTGAGGCGGCGGAGACGCGGCGGCTCAGGCGGCGGCGGCGGCGG 1399
QY 901 ACAGAGGAGGAGCTTCTGCGGCGGCGGCTTTTGGCGGCGG 941
DB 1400 ACAGAGGAGGAGCTTCTGCGGCGGCGGCTTTTGGCGGCGG 1440
```

Db 860 GGGTGGAGGAGCTCTCAAGGAIITCTGAGCCGATGGCAGGAGGAGGCGCAAGGCA 919
QY 421 AGAGGGCGCGGAGCAACACCTTGAACCTTCGGGGCGCGCTCCGGGGCCCGCTCGCC 480
Db 920 AGAGGGCGCGGAGCAACACCTTGAACCTTCGGGGCGCGCTCCGGGGCCCGCTCGCC 379
QY 481 AGCACTCTCCCTACGGCGGCTCGGGCCCGCGGCGCACCGCGGCTCTGTCGCGCCCGCCCTCT 540
Db 980 AGCACTCTCCCTACGGCGGCTCGGGCCCGCGGCGCACCGCGGCTCTGTCGCGCCCGCCCTCT 1039
QY 541 CCGTAGCGCGGAGGAGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 1040 CCGTAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
QY 601 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 560
Db 1100 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1159
QY 661 AAGTTTAAATGCGCGCGCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Db 1160 AAGTTTAAATGCGCGCGCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1219
QY 721 CGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 1220 CGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1279
QY 781 CTCGGGGCTCGGGTCTACGGCGGCGAGCGGGCGCTCTCGGGCGGCGGCGGCGGAGT 840
Db 1280 CTCGGGGCTCGGGTCTACGGCGGCGAGCGGGCGCTCTCGGGCGGCGGCGGCGGAGT 1339
QY 841 GGTGTCTCTGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 1340 GGTGTCTCTGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1399
QY 901 ACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 941
Db 1400 ACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

RESULT 3
AL139005/c
LOCUS Human DNA sequence from clone RP11-95G6 on chromosome 13, complete
DEFINITION sequence.
ACCESSION AL139005
VERSION AL139005.12 GI:14018246
KEYWORDS Hig.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 102119)
Direct Submission
Submitted (09-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerquest@sanger.ac.uk
On May 11, 2001 this sequence version replaced gi:13274245.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations.
Together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Emi, EMBL; Swi,

SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr13
RP11-95G6 is from the library RP11-11.1 constructed by the group of
Piet de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.5
IMPORTANT: This sequence is not the entire insert of clone
RP11-95G6. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-95G6 is at 102119 in this
sequence. The true left end of clone RP11-57H24 is at 59065 in this
sequence. The true right end of clone RP11-502P18 is at 100 in this
sequence.

FEATURES
source location/Qualifiers
1..102119
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-95G6"
/clone_lib="RP11-11.1"
851..979
repeat_region
/note="FLAM_A repeat: matches 1..129 of consensus"
repeat_region
1038..1087
/note="25 copies 2 mer gt 92% conserved"
repeat_region
1044..1087
/note="11 copies 4 mer ggt 100% conserved"
repeat_region
2014..2102
/note="12 repeat: matches 2645..2742 of consensus"
repeat_region
3151..3398
/note="MIR repeat: matches 2..252 of consensus"
repeat_region
3745..3813
/note="3 copies 23 mer 82% conserved"
repeat_region
4238..4480
/note="AluSq repeat: matches 1..243 of consensus"
repeat_region
4481..4512
/note="8 copies 4 mer aaat 93% conserved"
repeat_region
5434..5467
/note="17 copies 2 mer aa 82% conserved"
repeat_region
6041..6085
/note="9 copies 5 mer ttatg 84% conserved"
repeat_region
6084..6387
/note="AluSx repeat: matches 1..300 of consensus"
repeat_region
7056..7367
/note="AluSg repeat: matches 1..310 of consensus"
repeat_region
8032..8343
/note="AluSx repeat: matches 1..312 of consensus"
repeat_region
8585..8771
/note="MER5B repeat: matches 84..178 of consensus"
repeat_region
9322..9957
/note="12 copies 3 mer tta 94% conserved"
repeat_region
9358..10079
/note="AluJo/FRAM repeat: matches 156..282 of consensus"
repeat_region
10539..10707
/note="L1M25 repeat: matches 7772..7958 of consensus"
repeat_region
10739..11016
/note="L1M24 repeat: matches 7645..7820 of consensus"
repeat_region
12636..12928
/note="AluB repeat: matches 1..307 of consensus"
repeat_region
13961..14177
/note="MIR repeat: matches 46..252 of consensus"
repeat_region
14564..14862
/note="AluSx repeat: matches 1..305 of consensus"
repeat_region
15571..15708
/note="6 copies 23 mer 70% conserved"
repeat_region
15576..15701
/note="63 copies 2 mer tt 71% conserved"
repeat_region
15890..15898
/note="7 copies 17 mer 74% conserved"

```
repeat_region 15525..15704
/Note="20 copies 4 mer tctt 87% conserved"
repeat_region 15708..16014
/Note="AluX repeat: matches 1..308 of consensus"
repeat_region 16022..16126
/Note="AluSg/x repeat: matches 202..306 of consensus"
repeat_region 17111..17199
/Note="MIR repeat: matches 1..94 of consensus"
repeat_region 17832..17897
/Note="MIR repeat: matches 193..260 of consensus"
repeat_region 17845..18023
/Note="L2 repeat: matches 2567..2739 of consensus"
repeat_region 18059..18137
/Note="3 copies 23 mer 87% conserved"
repeat_region 18494..18777
/Note="AluSp repeat: matches 3..285 of consensus"
repeat_region 19334..19385
/Note="L2 repeat: matches 2648..2701 of consensus"
repeat_region 19350..19582
/Note="MIR repeat: matches 2..241 of consensus"
repeat_region 19810..19971
/Note="MIR repeat: matches 85..258 of consensus"
repeat_region 20842..20921
/Note="MIR repeat: matches 54..140 of consensus"
repeat_region 20566..21034
/Note="FLAM_A repeat: matches 51..128 of consensus"
repeat_region 21009..21072
/Note="AluA8 repeat: matches 134..210 of consensus"
repeat_region 21868..22166
/Note="AluB repeat: matches 1..302 of consensus"
repeat_region 22892..23190
/Note="AluY repeat: matches 1..299 of consensus"
repeat_region 25148..25236
/Note="MIR repeat: matches 52..147 of consensus"
repeat_region 27232..27532
/Note="AluSg repeat: matches 1..313 of consensus"
repeat_region 29375..29442
/Note="34 copies 2 mer tt 66% conserved"
repeat_region 29702..30027
/Note="L2 repeat: matches 2356..2710 of consensus"
repeat_region 31857..31999
/Note="MIR repeat: matches 25..141 of consensus"
repeat_region 35931..36016
/Note="MER3 repeat: matches 1..208 of consensus"
repeat_region 37895..38193
/Note="AluSg repeat: matches 1..299 of consensus"
repeat_region 40064..40284
/Note="MIR repeat: matches 14..241 of consensus"
repeat_region 40693..40743
/Note="L2 repeat: matches 2647..2699 of consensus"
repeat_region 42794..43196
/Note="MIR repeat: matches 3..420 of consensus"
repeat_region 43395..43657
/Note="MIR repeat: matches 77..140 of consensus"
repeat_region 43937..44142
/Note="MIR repeat: matches 14..252 of consensus"
repeat_region 44173..44831
/Note="MER50 repeat: matches 1..709 of consensus"
repeat_region 44554..45088
/Note="FLAM_C repeat: matches 2..129 of consensus"
repeat_region 45992..46035
/Note="22 copies 2 mer ac 90% conserved"
repeat_region 45996..46035
/Note="10 copies 4 mer acac 95% conserved"
repeat_region 47173..47453
/Note="LIMB8 repeat: matches 2311..2594 of consensus"
repeat_region 47454..47743
/Note="AluX repeat: matches 1..292 of consensus"
repeat_region 47744..48488
/Note="LIMB8 repeat: matches 2594..3347 of consensus"
repeat_region 48489..48761
/Note="AluX repeat: matches 1..275 of consensus"
repeat_region 48762..50491
```

```
repeat_region 50492..50780
/Note="LIMB8 repeat: matches 3347..5028 of consensus"
repeat_region 50781..51024
/Note="AluB repeat: matches 1..285 of consensus"
repeat_region 51044..51179
/Note="LIMB8 repeat: matches 5028..5274 of consensus"
repeat_region 51190..52061
/Note="8 copies 17 mer 66% conserved"
repeat_region 52293..52577
/Note="LIMB8 repeat: matches 5271..6171 of consensus"
repeat_region 53786..53900
/Note="AluJ repeat: matches 1..295 of consensus"
repeat_region 53786..53875
/Note="5 copies 23 mer 68% conserved"
repeat_region 54044..54138
/Note="22 copies 4 mer agga 68% conserved"
repeat_region 54435..54742
/Note="MIR repeat: matches 46..142 of consensus"
repeat_region 54900..54983
/Note="AluX repeat: matches 1..308 of consensus"
repeat_region 55026..55061
/Note="42 copies 2 mer gg 64% conserved"

Query Match 99.8%; Score 939.4; DB 9; Length 1021119;
Best Local Similarity 99.9%; Pred. No. 7.8e-144;
Matches 940; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGACGTTCCTCCAAAG 60
Db 99118 GCCGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGACGTTCCTCCAAAG 99059

QY 51 GTCCCAACAGCGTCAGTTCCTTCAGGCTCCAGTTCAGTTCAGTTCGTCGAGGT 120
Db 99058 GTCCCAACAGCGTCAGTTCCTTCAGGCTCCAGTTCAGTTCAGTTCGTCGAGGT 98999

QY 121 CTCGGTTCCTTCAGTTCCTTCGGGACAGTCTGAAGGGTCAGGAGCGGGGACAGC 180
Db 98998 CTCGGTTCCTTCAGTTCCTTCGGGACAGTCTGAAGGGTCAGGAGCGGGGACAGC 98939

QY 181 GCGGAAGACAGGACAGGAGGACAGCGGACCTCCGCTTCAGTTCCTTCGTCGACGAC 240
Db 98938 GCGGAAGACAGGACAGGAGGACAGCGGACCTCCGCTTCAGTTCCTTCGTCGACGAC 98879

QY 241 ACCGTTCGGGAGGCGCGCCAGCTTCCTTCGATCGGACTTCGCGCCCTAGGCGCAGGC 300
Db 98878 ACCGTTCGGGAGGCGCGCCAGCTTCCTTCGATCGGACTTCGCGCCCTAGGCGCAGGC 98819

QY 301 GCGGAGCTTCAGCTTCCTTCCTCCAGTTCGGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 98818 GCGGAGCTTCAGCTTCCTTCCTCCAGTTCGGGCGGCGGCGGCGGCGGCGGCGGCGG 98759

QY 361 GCGTGGAGGAGTTCGAAGGATTCCTCAGCGCGATGGCAGGAGGAGGCGCAAGGCA 420
Db 98758 GCGTGGAGGAGTTCGAAGGATTCCTCAGCGCGATGGCAGGAGGAGGCGCAAGGCA 98699

QY 421 AGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Db 98698 AGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 98639

QY 481 AGACCTTCCTCCAGGCGGCTTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 540
Db 98638 AGACCTTCCTCCAGGCGGCTTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 98579

QY 541 CCGTACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 98578 CCGTACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 98519

QY 501 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 98518 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 98459

QY 561 AGGTTATTAATCGCCCGGCTTCGCTTCATCGAGGTCGCGGGAGGAGGAGGAGGAGGAGG 720
Db 98459 AGGTTATTAATCGCCCGGCTTCGCTTCATCGAGGTCGCGGGAGGAGGAGGAGGAGGAGG 98400
```


Db	98458	AAAGTTATAAAATGCCCCCGCCCTCGGCTGCTCTTCAAGAGTCCGGGGAGGCTCGA	98399
QY	721	GGCGCCAGCGGACACTCTCTGGCTCTCTCCCGGAGCGGCGGCTCGAGCGG	980
Db	98398	GGCGCCAGCGGACACTCTCTGGCTCTCTCCCGGAGCGGCGGCTCGAGCGG	98339
QY	781	CTCCGGGCTCGGCTGAGCGGCGGAGCGGCTCTGGCGGAGGATACCGGGGAGT	840
Db	98338	CTCCGGGCTCGGCTGAGCGGCGGAGCGGCTCTGGCGGAGGATACCGGGGAGT	98279
QY	841	GGTGTCTCTGGCTGAGCGGCGGAGCGGCTCTAGGCGGCGGCGGCGGCGG	900
Db	98278	GGTGTCTCTGGCTGAGCGGCGGAGCGGCTCTAGGCGGCGGCGGCGGCGG	98219
QY	901	ACGAGGAGCGGACTCTCGCGCGCGGCTCTTGGCGCGGCGG 941	
Db	98218	ACGAGGAGCGGACTCTCGCGCGCGGCTCTTGGCGCGGCGG 98178	
SULT 4			
3168			
LOCUS			
DEFINITION	E10168	840 bp	DNA
ACCESSION	E10168	Promoter of human flt gene encoding receptor type tyrosine kinase.	PAT 23-SEP-1997
VERSION	E10168.1	GI:22025997	
KEYWORDS	JP 1995289263-A/1.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
TITLE	Shibuya M.		
JOURNAL	PROMOTER FOR GENE MANIFESTATION		
COMMENT	Patent: JP 1995289263-A 1 07-NOV-1995;		
	DAI ICHI SEIYAKU CO LTD		
	OS Homo sapiens (human)		
	PN JP 1995289263-A/1		
	FD 07-NOV-1995		
	PF 22-APR-1994		
	PI SHIBUYA MASASHI		
	PC C12N15/09,A61K31/70,A61K45/00,C12N5/10,C12P21/02//C12P21/04,		
	PC C12N1/68		
	PC (C12N5/10,C12P1:91),(C12P21/02,C12P1:91);		
	CC strandedness: Double;		
	CC topology: Linear;		
	CC hypothetical: No;		
	CC anti-sense: No;		
Key	Location/Qualifiers		
FI	Source	1..840	
FI	/organism='Homo sapiens'		
FI	/tissue_type='placenta'		
FI	promoter	1<..489	
FI	TATA_signal	457..462	
FI	CDS	771..3840	
FI	/product='receptor type tyrosine kinase'		
FI	misc_feature	343..347	
FI	/note='Ets core motif'		
FI	misc_feature	350..354	
FI	/note='Ets core motif'		
FI	misc_feature	435..440	
FI	/note='Ets core motif'		
FEATURES	Location/Qualifiers		
source	1..840		
	/organism='Homo sapiens'		
	/db_xref='taxon:9606'		
BASE COUNT	120 a	269 c	336 g
ORIGIN	115 t		
Query Match	70.9%	Score 567;	DB 5; Length 840;
Best Local Similarity	98.6%	Fred. No. 4.9e-99;	
Matches 725; Conservative	0;	Mismatches 5;	Indels 5; Gaps 5;

QY	208	CCGAGCTCGGCTCAGTCTCTGTCGCGAGACACCGTTCGGAGGCGCGGCGGCTTCC	267
Db	1	CCGAGCTCGGCTCAGTCTCTGTCGCGAGACACCGTTCGGAGGCGCGGCGGCTTCC	60
QY	258	CTTGGATCGGACTTTCGGCCCTAGGGCAGCGGCGGAGCTTACGCTTGTCCCTTCC	327
Db	61	CTTGGATCGGACTTTCGGCCCTAGGGCAGCGGCGGAG-TTACGCTTGTCTT-CC	118
QY	328	CAGTTTCGGGCGGCGGCGGCGGCTGAGTAAGCGGCTGGAGGAGTCTCCAGGATTTCC	387
Db	119	CAGTTTCGGGCGGCGGCGGCGGCTGAGTAAGCGGCTGGAGGAGTCTCCAGGATTTCC	178
QY	388	TGAGCGCATGGCGAGGAGGCGGCAAGGCGGAGGCGGCGGAGCAAGACCTGAAC	447
Db	179	TGAGCGCATGGCGAGGAGGCGGCAAGGCGGAGGCGGCGGAG-CAAGACCTGAAC	237
QY	448	CTCCGCGGCGGCGGCTCCCGGCGGCGGCTGCGCACCTCCCGCGGCTCGGCGG	507
Db	238	CTCCGCGGCGGCGGCTCCCGGCGGCGGCTGCGCACCTCCCGCGGCTCGGCGG	297
QY	508	GGGCGCATCGGCGGCTCGTCCGCGGCGGCGGCTCGTCCGTAAGCGGAGGAGG-CAAGCCTGG	565
Db	298	GGGCGCATCGGCGGCGGCTCGTCCGCGGCGGCGGCTCGTCCGTAAGCGGAGGAGG-CAAGCCTGG	357
QY	567	GAGCAAGAGAGGAGTGGTGGGAGGCGGATGAGGGGTGGGGACCCCTTGACGTACCA	625
Db	358	GAGCAAGAGAGGAGTGGTGGGAGGCGGATGAGGGGTGGGGACCCCTTGACGTACCA	417
QY	627	GAGGAGGTGGCGGGGTAGGAGTGGGCTGGGGAAGGTATTAATCGGCGGCGGCTCG	685
Db	418	GAGGAGGTGGCGGGGTAGGAGTGGGCTGGGGAAGGTATTAATCGGCGGCGGCTCG	477
QY	587	GCTGCTCTTCATCGAGGTCCCGGAGGCTCGGAGGCGGCGGCGGAGGACACTCTCTGG	745
Db	478	GCTGCTCTTCATCGAGGTCCCGGAGGCTCGGAGGCGGCGGCGGAGGACACTCTCTGG	537
QY	747	CTGCTCTCGGCGGCGGCGGCTCGGAGGCGGCTCGGAGGCGGCTCGGAGGCGGCGG	805
Db	538	CTGCTCTCGGCGGCGGCGGCTCGGAGGCGGCTCGGAGGCGGCTCGGAGGCGGCGG	597
QY	807	CGGCGGCTCGGCGGCGGAGTACCGGGGAGGTGGTGTCTCTGCTGAGCGGCGGAG	865
Db	598	CGAGCG-CTGGCGGCGGAGTACCGGGGAGGTGGTGTCTCTGCTGAGCGGCGGAG	655
QY	867	ACGGGCGCTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	925
Db	557	ACGGGCGCTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	925
QY	927	GCTTTGGCGGCGG 941	
Db	717	GTCGTTGGCGGCGG 731	
RESULT 5			
AX251105			
LOCUS			
DEFINITION	AX251105	6316 bp	DNA
ACCESSION	AX251105	Sequence 73 from Patent WO0168912.	linear
VERSION	AX251105.1	GI:15984528	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE	1 (bases 1 to 6316)		
AUTHORS	Olek A., Piepenbrock, C. and Berlin, K.		
TITLE	Diagnosis of diseases associated with tumor suppressor genes an;		
JOURNAL	Patent: WO 0168912-A 73 20-SEP-2001;		
FEATURES	Location/Qualifiers		
source	1..6316		
	/organism='synthetic construct'		
	/db_xref='taxon:32630'		

```

* RESULT 5
AX251868
LOCUS AX251868 6316 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 129 from Patent WO0168911.
ACCESSION AX251868
VERSION AX251868.1 GI:15985223
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE 1 (bases 1 to 6316)
AUTHORS Olek A. Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with the cell cycle
JOURNAL Patent: WO 0168911-A 129 20-SEP-2001;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
source
1..6316
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1490 a 255 c 1803 g 2768 t
ORIGIN
Query Watch 65.1%; Score 613; DB 6; Length 6316;
Best Local Similarity 78.2%; Pred. No. 1.7e-90;
Matches 736; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 1 GCCGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGAGCTCTCCAAAG 60
Db 4272 GTCGAAAAGATACGATACGTTTTTTGGGATTCGATTTGTTTAAAG 4331

QY 61 GTGCCAGCAGGCTCAGTTCCCTCAGGCTCCAGCTTCAGCTTCGCTGCGGAGGT 120
Db 4332 GTGTAAGTAAGCGGTAGTTTTTTAGCGGTTTGTAGTTTGTGCGAGGT 4391

QY 121 CTCGGTGCCTTCTAGACTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAC 180
Db 4392 TTTCGGTGTTCCTAGATTTTCGGGATTTTCGGGATTTTCAGGCGGCGGATGC 4451

QY 181 GCGGAGAGCAGCAGGAGGAGAGCGGAGCTTCAGCTTCAGCTTCGCTGCGGAGGCA 240
Db 4452 GCGGAGAGGATAGTAAAGGAGATTCGGAATTTTCGGTTCGTTTAAAGAT 4511

QY 241 ACCGTTCGGAGCGCGGCGGACCTTCCTTCGATCGACTTTCCGCGGCGGCGGAG 300
Db 4512 ATCGTCGGAGCGGCGGTAGTTTTTTTGGATCGGATTTTCGTTTAAAGGAT 4571

QY 301 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 360
Db 4572 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 4631

QY 361 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 420
Db 4632 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 4691

QY 421 ACAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Db 4692 ACAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4931

QY 481 AGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 4932 AGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5491

QY 541 CGGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 5492 CGGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6061

QY 601 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 6062 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6721

QY 661 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 6722 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7261

QY 721 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 7262 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7921

QY 781 CTCGGGCGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Db 7922 CTCGGGCGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8461

QY 841 GGTTCCTTCGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db 8462 GGTTCCTTCGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9061

QY 901 ACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 9062 ACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9661

QY 961 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Db 9662 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 10261

```

```

* RESULT 6
AX251868
LOCUS AX251868 6316 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 129 from Patent WO0168911.
ACCESSION AX251868
VERSION AX251868.1 GI:15985223
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE 1 (bases 1 to 6316)
AUTHORS Olek A. Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with the cell cycle
JOURNAL Patent: WO 0168911-A 129 20-SEP-2001;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
source
1..6316
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1490 a 255 c 1803 g 2768 t
ORIGIN
Query Watch 65.1%; Score 613; DB 6; Length 6316;
Best Local Similarity 78.2%; Pred. No. 1.7e-90;
Matches 736; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 1 GCCGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGAGCTCTCCAAAG 60
Db 4272 GTCGAAAAGATACGATACGTTTTTTGGGATTCGATTTGTTTAAAG 4331

QY 61 GTGCCAGCAGGCTCAGTTCCCTCAGGCTCCAGCTTCAGCTTCGCTGCGGAGGT 120
Db 4332 GTGTAAGTAAGCGGTAGTTTTTTAGCGGTTTGTAGTTTGTGCGAGGT 4391

QY 121 CTCGGTGCCTTCTAGACTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAC 180
Db 4392 TTTCGGTGTTCCTAGATTTTCGGGATTTTCGGGATTTTCAGGCGGCGGATGC 4451

QY 181 GCGGAGAGCAGCAGGAGGAGAGCGGAGCTTCAGCTTCAGCTTCGCTGCGGAGGCA 240
Db 4452 GCGGAGAGGATAGTAAAGGAGATTCGGAATTTTCGGTTCGTTTAAAGAT 4511

QY 241 ACCGTTCGGAGCGCGGCGGACCTTCCTTCGATCGACTTTCCGCGGCGGCGGAG 300
Db 4512 ATCGTCGGAGCGGCGGTAGTTTTTTTGGATCGGATTTTCGTTTAAAGGAT 4571

QY 301 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 360
Db 4572 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 4631

QY 361 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 420
Db 4632 GCGGAGCTTCAGCTTCCTGGGACAGTCTGAAGGGTTCAGGCGGCGGAGGCA 4691

QY 421 ACAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Db 4692 ACAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4931

QY 481 AGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 4932 AGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5491

QY 541 CGGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 5492 CGGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6061

QY 601 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 6062 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6721

QY 661 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 6722 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7261

QY 721 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 7262 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7921

QY 781 CTCGGGCGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Db 7922 CTCGGGCGGCTCGGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8461

QY 841 GGTTCCTTCGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db 8462 GGTTCCTTCGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9061

QY 901 ACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 9062 ACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9661

QY 961 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Db 9662 AGGTTATTAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 10261

```



```
QY 121 CTCGGTCCCTTCTAGACTTCTGGGACAGTCTGAAGGGTCAAGACGCGGGACAGC 180
Db 4392 TTTCGGTCTTTTATATTTTGGGATATTTGAAGGGTTAGAGCGCGGATAGC 4451
QY 181 CGGGAGACAGCAGGAGGAGACAGCGGAGTGGCGCTCAGTCTCGTCCGACCAAGC 240
Db 4452 CGGGAGAGAGTGAAGGGAGATAGTCGATGGGTTTATGTTTTCGTTTGAAGAT 4511
QY 241 ACCGTCCGGAGGCGGGCCAGCTTCCCTTGGATCGACTTTCCGCCCTTAGGCCAGGC 300
Db 4512 ATCGTCGGAGCGCGGTAGTTTTTTTGGATCGGATTTTCGTTTTTAGGGTAGCC 4571
QY 301 GCGGAGCTTACGCTTGTCCCTTCCCGAGTTTCGGGCGCCCGCCAGAGCTGAGTAGCC 360
Db 4572 GCGGAGTTTATGTTTGTTTTTTATGTTTCGGGCGGTTTTAGAGTTCAGTAGTC 4631
QY 361 GGTGAGGAGTCTGCAAGGATTTCCCTGAGCGCATGGGAGGAGGAGGCGCAAGGCA 420
Db 4632 GGTGAGGAGTTTGTAGGATTTTTCAGCGGATGGTAGGAGGAGGTTAAGGGA 4591
QY 421 AGAGGCGCGGAGCAAGACCTGAACTCGCGGGCGCGCTTCCGGGCGCGCTGCC 480
Db 4692 AGAGGCGCGGAGTAAAGATTTTGAATTTTCGGGGTTCGGGTTTCGGTCTGTT 4751
QY 481 ACCACTTCCCGACCTCGCTCGCGCGCGGACCTTCCCTTCTGCGCGCGCGCTCT 540
Db 4752 AGTATTTTTTACCGCGTTTCGGTTTCGGTTTTCGTTTTCGTTTTTTCGTTTT 4811
QY 541 CCGTAGCGGAGGAGTCTGAGGAGTGGGAGAGAGGAGGTTGGGAGGCGGATAG 600
Db 4812 TCGTAGTCTAGGAGGAGTTCAGCTTATAGAGGAGGTTTCGGGTTAGGAAGTTCG 4931
QY 601 GGTGGGAGCCCTTCACTTACAGAGAGAGGTTCGGGGTAGGAAGTGGCTGGGGA 660
Db 4872 GGTGGGAGTTTTCACGTTATAGAGGAGGTTTCGGGTTAGGAAGTTCGAGTT 4991
QY 661 AGGTTATATATCGCCCGCTCGCTTCTATCAGAGTTCGGGGAGGCTCGGA 720
Db 4932 AGGTTATATATCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTT 780
QY 721 GCGCGGAGGCGGACTCTCTCGCTCTCTCGCGGAGGCGGCGGCTCGGAGCGGG 780
Db 4992 GCGGTTAGCGGATATTTTTTTCGTTTTTTCGTTTTTTCGTTTTTTCGTTTT 840
QY 781 CTCGGGCTCGGTGACGCGGACGCGGCGCTGCGCGGAGGATACCGCGGAGT 840
Db 5052 TTTCGGGGTTCGGGTGATAGCGGTTTTCGGGCGGTTTTCGGGCGGAGGATAT 5111
QY 841 GGTGTCTCTGCTGAGCGGAGACGCGGCTCAGGCGGCGGCGGCGGCGGCGGA 900
Db 5112 GGTGTCTCTGCTGAGCGGAGACGCGGCTTTCGGGCGGTTTTCGGGCGGAGG 5171
QY 901 ACAGAGGAGGAGTCTGGGCGCGGCTCTTTTGGCGCGGG 941
Db 5172 ACAGAGGAGGAGTTTGGCGGTCGCTGTTTGGTTCGCGG 5212
```

RESULT 9
AX458547 6321 bp DNA linear PAT 08-JUL-2002

LOCUS
DEFINITION
Sequence 93 from Patent WO0245454.
ACCESSION
AX458547
VERSION
AX458547.1 GI:21725211

KEYWORDS
SOURCE
synthetic construct.
synthetic construct
artificial sequences.

REFERENCE
1
AUTHORS
Schacht, O.
TITLE
Diagnosis of diseases associated with angiogenesis
JOURNAL
Patent: WO 0245454-A 93 13-JUN-2002;
Epigenomics AG (DE)

FEATURES
Location/Qualifiers

```
source 1. 5321  
/organism="synthetic construct"  
/db_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 1489 a 258 c 1805 g 2769 t  
ORIGIN  
Query Match 55.1%; Score 613; DB 6; Length 6321;  
Best local similarity 78.2%; Pred. No. 1.7e-90;  
Matches 736; Conservative 0; Mismatches 205; Indels 0; Gaps 0;  
QY 1 GCCGAAAAGACAGCAGCAGCAGCTCCCTGGAGCAGTGGTGGTTCAGAGTCTCCCAAAG 60  
Db 4272 GTCCGAAAAGACAGCAGCAGCTTTTGGGATTTGATTTGATTTGATTTTAAAG 4331  
QY 61 GTCCCAAGCAGAGCTCAGTTCCCTCAGGCGCTCCAGTTCAGTCCCTTTGCCGAGGT 120  
Db 4332 GTGTAAGTAAGCTAGTTTTCAGGCGTTTGGTTTGGTTTGGTTTGGTTTGGTTT 4391  
QY 121 CTCGGTCCCTTCTAGACTTCTCGGACAGTCTGAAGGGTCAAGAGCGCGGACAGC 180  
Db 4392 TTTCGGTCTTTTATATTTTGGGATATTTGAAGGGTTAGAGCGCGGATAGC 4451  
QY 181 GCGGAAAGACAGCAGGAGGAGACAGCGGACTCCGCTCAGTCTCGCTCCCAAGAAC 240  
Db 4452 GCGGAAAGATAGGAGGAGATAGTCGATTTGATTTTTCGTTTGAAGAT 4511  
QY 241 ACCGTCCGGAGGCGGCGGAGTTCCTTGGATCGACTTTCCGCCCTTAGGCCAGGC 300  
Db 4512 ATCGTCGGAGCGCGGTAGTTTTTGGATCGGATTTTCGTTTTCGTTTTCGTTT 4571  
QY 301 GCGGAGCTTACGCTTGTCCCTTCCCGAGTTTCGGGCGCGCCAGAGCTCAGTAGCC 360  
Db 4572 GCGGAGTTTATGTTTGTTTTTTATGTTTCGGGCGGTTTTAGAGTTCAGTAGTC 4631  
QY 361 GGTGAGGAGTCTCAGAGATTTCTTACGCGGATGGGAGAGGAGGCGCAAGGCA 420  
Db 4632 GGTGAGGAGTTTGAAGATTTTTCAGCGCATGGTAGGAGGAGGTTAAGGGA 4591  
QY 421 AGAGGCGCGGAGCAAGACCTGAACTCGCGGGCGCGCTCCCGCGCGCGCTGCC 480  
Db 4692 AGAGGCGCGGAGTAAAGATTTTGAATTTTCGCGGTTTCGTTTTCGTTTTCGTT 4751  
QY 481 AGCACTCTCCCGACGCTCGCGCGCGGCGGCGGCTCGCTCGCGCGCGCGCTCT 540  
Db 4752 AGTATTTTTTACCGCGTTTCGGTTTCGGTTTATGTTTTCGTTTTCGTTTTT 4811  
QY 541 CCGTAGCGGAGGAGCAGCAGCTCGGAGGAGAGAGGTTAGTGGGAGCGGATAG 600  
Db 4812 TCGTAGTCTAGGAGGAGTTCAGCTTATAGAGGAGGTTTCGGGTTAGGAAGTTCG 4931  
QY 601 GGTGGGAGCCCTTCACTTACAGAGAGAGGTTCGGGGTAGGAAGTGGCTGGGGA 660  
Db 4872 GGTGGGAGTTTTCACGTTATAGAGGAGGTTTCGGGTTAGGAAGTTCGAGTT 4991  
QY 721 GCGCGGAGGCGGACTCTCTCGCTCTCTCGCGGAGGCGGCGGCTCGGAGCGGG 780  
Db 4992 GCGGTTAGCGGATATTTTTTTCGTTTTTTCGTTTTTTCGTTTTTTCGTTTT 840  
QY 781 CTCGGGCTCGGTGACGCGGACGCGGCGCTGCGCGGAGGATACCGCGGAGT 840  
Db 5052 TTTCGGGGTTCGGGTGATAGCGGTTTTCGGGCGGTTTTCGGGCGGAGGATAT 5111  
QY 841 GGTGTCTCTGCTGAGCGGAGACGCGGCTCAGGCGGCGGCGGCGGCGGCGGA 900  
Db 5112 GGTGTCTCTGCTGAGCGGAGACGCGGCTTTCGGGCGGTTTTCGGGCGGAGG 5171  
QY 901 ACAGAGGAGGAGTCTGGGCGCGGCTCTTTTGGCGCGGG 941  
Db 5172 ACAGAGGAGGAGTTTGGCGGTCGCTGTTTGGTTCGCGG 5212
```

Db 30630 TCITGTCCTATGCCAGA 30614

RESULT 3

AC128405/c

LOCUS

DEFINITION

AC128409

AC128409.1 GI:21905078

HIG: HTGS-PHASE1

SOURCE

ORGANISM

Rattus norvegicus

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 144090)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaralungu,H.C., Are,J.R., Ayele,M., Banks,J.,

Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnia,D.,

Bouck,J., Bowler,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhal,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Correll,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Litcharge,O., Lieu,C., Liu,J., Liu,M., Louissegh,H.,

Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,

Miner,G., Miner,Z., Mitchell,T., Monabbat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,

Oraquye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojibokan,I., Roife,M., Ruiz,S., Savary,G.,

Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,

Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,

Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,

Wu,C., Wu,T., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 144090)

Worley,K.C.

Direct Submission

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYFL

Center clone name: CH230-448N2

Summary Statistics

Sequencing Vector: Plasmid:

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly: Phrap; version 0.950325

Consensus quality: 101711 bases at least Q40

Consensus quality: 106747 bases at least Q30

Consensus quality: 110016 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 57 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1636: contig of 1636 bp in length

* 1637 1736: gap of unknown length

* 1737 3403: contig of 1667 bp in length

* 3404 3503: gap of unknown length

* 4554: contig of 1051 bp in length

* 4555 4654: gap of unknown length

* 4655 5864: contig of 1210 bp in length

* 5865 5964: gap of unknown length

* 5965 7113: contig of 1151 bp in length

* 7116 7213: gap of unknown length

* 8510: contig of 1295 bp in length

* 8511 8610: gap of unknown length

* 8611 9955: contig of 1345 bp in length

* 9256 10055: gap of unknown length

* 11380: contig of 1325 bp in length

* 11480: gap of unknown length

* 11481 12607: contig of 1127 bp in length

* 12608 13653: contig of 1156 bp in length

* 13664 13963: gap of unknown length

* 13964 13963: contig of 1399 bp in length

* 13963 14621: gap of unknown length

* 14621 17024: contig of 1562 bp in length

* 17025 17124: gap of unknown length

* 17125 18879: contig of 1755 bp in length

* 18880 18979: gap of unknown length

* 18980 20913: contig of 1934 bp in length

* 20914 21013: gap of unknown length

* 21014 2216: contig of 2203 bp in length

* 2217 2316: gap of unknown length

* 2317 2554: contig of 2238 bp in length

* 2555 2654: gap of unknown length

* 2655 27062: contig of 1408 bp in length

* 27063 27162: gap of unknown length

* 27163 28572: contig of 1410 bp in length

* 28573 28672: gap of unknown length

* 28673 29930: contig of 1258 bp in length

* 29931 30030: gap of unknown length

* 30031 32527: contig of 2457 bp in length

* 32528 32627: gap of unknown length

* 32628 33881: contig of 1254 bp in length

* 33882 33981: gap of unknown length

* 33982 3613: contig of 2632 bp in length

* 3614 36713: gap of unknown length

* 36714 38530: contig of 1817 bp in length

* 38531 38630: gap of unknown length

* 38631 40221: contig of 1591 bp in length

* 40222 40321: gap of unknown length

* 40322 42216: contig of 1895 bp in length

* 42217 42316: gap of unknown length

* 42317 44902: contig of 2586 bp in length

* 44903 45002: gap of unknown length

* 45003 47520: contig of 2518 bp in length

* 47521 47620: gap of unknown length

* 47621 49384: contig of 1764 bp in length


```

gene      1. .4017
/ gene="FLT1"
CDS       1. .4017
/ gene="FLT1"
/ note="VEGFR; tyrosine kinase"
/ product="vascular endothelial growth factor receptor"
/ protein_id="AAC16449.1"
/ db_xref="GI:3132831"
/ translation="MVSMDTGVLLCALLSLLIGSSSSKSLKDPSELSKIQHIMQ
AGTLLHOCRAEAAHKSLEPMVSKESERLSITSACGRNGKQFCSTLILTAQAHHT
GYSCYLAIVPTSKKTESAIYIFISDTGRFFVEMSEIPIHMTGEIRELVTPCRV
TSPNITVLAKEFPDITLIPDKARILWDSKGFILSNATYKEIGLLTCEATVNGHLKYT
NLTHTQNTIIVQISITPPVKLLRGHTLVNCTATPLNIRVQMTSPDEKAKRA
SVRRIDGNSHANIFYSLTIDKQNDKGLYICRVSGSPKSVNTSVHYDKAFI
TVKHKQVLTAVAGKSYLSMKVAPSPVWVLKGLPATEKSAVLYTRGSLII
KQVTEAGNVTILLTSKQSNVFNKLTATLVNKKPOIYKAVSPSPALYPLGSSQ
ILTCYAGIPOPTIKMFWPCNHSEAKDCFCNNEESILFADSNMGNRIESTQR
MAIEGKKMASTLVADSRISGIVICIASKNKVGTVGRNISFYITDVPNGFHNLEK
PTEGEDKLSCTVANKFLYRDVTWILLRTVNNRIMHYSISKMAITKEHSITLNIIM
NVSLODSTYACRARNVYICETLOKKEITIDQEAAPYLLRNSDHYVALSSSTLDC
HANGVEPQITWFKNNHIOQEPGIIILGPGSTLFIERVTEDEGVYHCKAIKQGSV
ESSAYLTVOQSDSKSNLELITCTCTVAAITLWILLIRKMKRSSSEIKIDYLSII
MDPDEVPLDEOCERLPYDASKNEFARERKLGKSLGRGAGKVVQASAFGIKASPTCR
TVAVKMLEGATASEYKALMTELKILTHIGHLVNALLGACTKQGGPLMVIYECYK
GNLSYLSKSRDLFFLNKDAALHMEPKKEPGLQEGKPRDLSVTSSEFASGFO
EDKLSUVEEEDSDGYKEPITMEDLISYFQVARGMEFLSRKCIHRDLAARNILL
SENNVYKICDFGLARDIYNPDYVRKGTIRPLAWMAPESIFDKIYKSDVNSYGVIL
LNEIFSLGSSPYPCVOMDEDFCSLRREGMRAPESYTPYIOMLDCWHRPKRP
FAELVEKGLDLOANVQDGKDIYPINALITGNSGFITSPAFSEDFKESISAPFN
SCSSDDVYVNAFKENSLERIKTFEELLNPATSMFDYQDSDSTILLASPLMKRFETWD
SAPKASLAKIDLVTSKESGLSDVSRSPFCHSSCGHVSEGRKRRFTYDHAELERKIA
CSPPDYNSVWLYSTPPI"
BASE COUNT 1236 a 915 c 503 g 963 t
ORIGIN

Query Match      38.0%; Score 114.4; DB 9; Length 4017;
Best Local Similarity 99.1%; Pred. No. 3.5e-23;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 66 AGACTGCACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 125
Db 3173 ATATCGACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 3232
QY 126 CCAAGACGACGCTGTGGTCTTACGAGATGATGCTGTGGAAAATCTTCCCTAGGT 181
Db 3233 CCAAGACGACGCTGTGGTCTTACGAGATGATGCTGTGGAAAATCTTCCCTAGGT 3288

RESULT 6
HSLT
LOCUS      HSLT      7680 bp      mRNA      linear      PRI 15-NOV-1993
DEFINITION Human flt mRNA for receptor-related tyrosine kinase.
ACCESSION  X51602
VERSION    X51602.1 GI:31431
KEYWORDS   flt gene; fms-related tyrosine kinase gene; tyrosine kinase.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 7680)
            Shibusawa, M.
            Direct Submission
            Submitted (02-JAN-1989) Shibusawa M., Institute of Medical Science,
            University of Tokyo, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108,
            Japan
            2 (bases 1 to 7680)
            Shibusawa, M., Yamaguchi, S., Yamane, A., Ikeda, T., Tojo, A.,
            Matsushima, H. and Sato, M.
            Nucleotide sequence and expression of a novel human receptor-type
            tyrosine kinase gene (flt) closely related to the fms family
            Oncogene 5 (4), 519-524 (1990)
            90221591

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE

PUBMED 2158033
REFERENCE 3 (bases 1 to 7680)
AUTHORS Han, H.-J., Fujiwara, I., Shin, S. and Nakamura, Y.
TITLE Dinucleotide repeat polymorphism in the 3' non-coding region of the
        FLT1 gene
JOURNAL Hum. Mol. Genet. 2 (12), 2204 (1993)
MEDLINE 94154724
PUBMED 8111405
COMMENT Data kindly reviewed (20-JUL-1990) by Shibusawa M.
FEATURES
        location/Qualifiers
        1. 7680
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="13"
        /map="13q12"
        /clone="3-7, 3-5"
        /tissue_type="placenta"
        230..4266
        /note="flt gene product (AA 1-1336)"
        /codon_start=1
        /protein_id="CAA35946.1"
        /db_xref="GI:31432"
        /db_xref="SWISS-PROT:P17948"
        /translation="MVSMDTGVLLCALLSLLIGSSSSKSLKDPSELSKIQHIMQ
        AGTLLHOCRAEAAHKSLEPMVSKESERLSITSACGRNGKQFCSTLILTAQAHHT
        GYSCYLAIVPTSKKTESAIYIFISDTGRFFVEMSEIPIHMTGEIRELVTPCRV
        TSPNITVLAKEFPDITLIPDKARILWDSKGFILSNATYKEIGLLTCEATVNGHLKYT
        NLTHTQNTIIVQISITPPVKLLRGHTLVNCTATPLNIRVQMTSPDEKAKRA
        SVRRIDGNSHANIFYSLTIDKQNDKGLYICRVSGSPKSVNTSVHYDKAFI
        TVKHKQVLTAVAGKSYLSMKVAPSPVWVLKGLPATEKSAVLYTRGSLII
        KQVTEAGNVTILLTSKQSNVFNKLTATLVNKKPOIYKAVSPSPALYPLGSSQ
        ILTCYAGIPOPTIKMFWPCNHSEAKDCFCNNEESILFADSNMGNRIESTQR
        MAIEGKKMASTLVADSRISGIVICIASKNKVGTVGRNISFYITDVPNGFHNLEK
        PTEGEDKLSCTVANKFLYRDVTWILLRTVNNRIMHYSISKMAITKEHSITLNIIM
        NVSLODSTYACRARNVYICETLOKKEITIDQEAAPYLLRNSDHYVALSSSTLDC
        HANGVEPQITWFKNNHIOQEPGIIILGPGSTLFIERVTEDEGVYHCKAIKQGSV
        ESSAYLTVOQSDSKSNLELITCTCTVAAITLWILLIRKMKRSSSEIKIDYLSII
        MDPDEVPLDEOCERLPYDASKNEFARERKLGKSLGRGAGKVVQASAFGIKASPTCR
        TVAVKMLEGATASEYKALMTELKILTHIGHLVNALLGACTKQGGPLMVIYECYK
        GNLSYLSKSRDLFFLNKDAALHMEPKKEPGLQEGKPRDLSVTSSEFASGFO
        EDKLSUVEEEDSDGYKEPITMEDLISYFQVARGMEFLSRKCIHRDLAARNILL
        SENNVYKICDFGLARDIYNPDYVRKGTIRPLAWMAPESIFDKIYKSDVNSYGVIL
        LNEIFSLGSSPYPCVOMDEDFCSLRREGMRAPESYTPYIOMLDCWHRPKRP
        FAELVEKGLDLOANVQDGKDIYPINALITGNSGFITSPAFSEDFKESISAPFN
        SCSSDDVYVNAFKENSLERIKTFEELLNPATSMFDYQDSDSTILLASPLMKRFETWD
        SAPKASLAKIDLVTSKESGLSDVSRSPFCHSSCGHVSEGRKRRFTYDHAELERKIA
        CSPPDYNSVWLYSTPPI"
        2279 a 1661 c 1739 g 2001 t
        BASE COUNT
        ORIGIN

Query Match      37.5%; Score 112.8; DB 9; Length 7680;
Best Local Similarity 98.3%; Pred. No. 1e-22;
Matches 114; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 66 AGACTGCACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 125
Db 3422 ATATCGACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 3481
QY 126 CCAAGACGACGCTGTGGTCTTACGAGATGATGCTGTGGAAAATCTTCCCTAGGT 181
Db 3482 CCAAGACGACGCTGTGGTCTTACGAGATGATGCTGTGGAAAATCTTCCCTAGGT 3537

RESULT 7
HSLT
LOCUS      HSLT      220965 bp      DNA      linear      HTG 15-MAY-2002
DEFINITION Homo sapiens chromosome 3 clone RP6-329g3 map 3p21.3, ***
ACCESSION  AJ312688
VERSION    AJ312688.2 GI:13559235
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE     human.
ORGANISM   Homo sapiens
            Homo sapiens chromosome 3 clone RP6-329g3 map 3p21.3, ***
            SEQUENCING IN PROGRESS ***; 26 ordered pieces.
            AJ312688
            HTG; HTGS_PHASE2
            human
            SOURCE
            ORGANISM
            Homo sapiens

```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1
Kiss, H., Yang, Y., Kiss, C., Andersson, K., Klein, G., Imreh, S. and
Dumanski, J. P.
The transcriptional map of the common eliminated region 1 (CICER1)
in 3p21.3
Eur. J. Hum. Genet. 10 (1), 52-61 (2002)
21906202
PUBMED
11896456
REFERENCE
2 (bases 1 to 220965)
Kiss, H.
Direct Submission
Submitted (01-APR-2001) Kiss H., Microbiology and Immunobiology
Center (MTC), Karolinska Institute, Box 280, Stockholm, S-17177,
SWEDEN
On Apr 5, 2001 this sequence version replaced gi:11954653.
The sequence is a consensus sequence of clone RP6-787C23 (1-140400
bp),
clone RP6-32923 (31212-220965 bp), clone RP6-14661 (partially,
1-6800 bp)
and clone RP6-18911 (partially, 1-108303 bp). The sequencing
contigs are
in order and the gaps between them are represented by 100 Ns.
Contig 1:
1-11731 bp Contig 2: 11832-26218 bp Contig 3: 26319-28347 bp
Contig 4:
28448-42160 bp Contig 5: 42261-55059 bp Contig 6: 55160-61578 bp
Contig 7: 61679-97342 bp Contig 8: 97443-117655 bp Contig 9:
117756-118727 bp
Contig 10: 118828-121834 bp Contig 11: 121935-127855 bp Contig
12:
127956-129383 bp Contig 13: 129484-131747 bp Contig 14:
131848-132316 bp
Contig 15: 132417-134455 bp Contig 16: 134556-135527 bp Contig
17:
135628-189051 bp Contig 18: 189152-189476 bp Contig 19:
189577-191375 bp
Contig 20: 191476-201473 bp Contig 21: 201574-202307 bp Contig
22:
202408-204878 bp Contig 23: 204979-213531 bp Contig 24:
213632-218109 bp
Contig 25: 218210-219800 bp Contig 26: 219901-220965 bp.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 11731: contig of 11731 bp in length
* 11732 11831: gap of 100 bp
* 11832 26218: contig of 14367 bp in length
* 26219 26318: gap of 100 bp
* 26319 28347: contig of 2029 bp in length
* 28348 28447: gap of 100 bp
* 28448 42160: contig of 13713 bp in length
* 42161 42360: gap of 200 bp
* 42361 55059: contig of 12699 bp in length
* 55060 55159: gap of 100 bp
* 55160 61578: contig of 6419 bp in length
* 61579 61678: gap of 100 bp
* 61679 97342: contig of 35664 bp in length
* 97343 97442: gap of 100 bp
* 97443 117655: contig of 20213 bp in length
* 117656 117755: gap of 100 bp
* 117756 118727: contig of 972 bp in length
* 118728 118827: gap of 100 bp
* 118828 121834: contig of 3007 bp in length
* 121835 121934: gap of 100 bp

* 121935 127855: contig of 5921 bp in length
* 127856 127955: gap of 100 bp
* 127956 129383: contig of 1428 bp in length
* 129384 129483: gap of 100 bp
* 129484 131747: contig of 2264 bp in length
* 131748 131847: gap of 100 bp
* 131848 132316: contig of 469 bp in length
* 132317 132416: gap of 100 bp
* 132417 134455: contig of 2039 bp in length
* 134456 134555: gap of 100 bp
* 134556 135527: contig of 972 bp in length
* 135528 135627: gap of 100 bp
* 135628 189051: contig of 53424 bp in length
* 189052 189151: gap of 100 bp
* 189152 189476: contig of 325 bp in length
* 189477 189576: gap of 100 bp
* 189577 191375: contig of 1799 bp in length
* 191376 191475: gap of 100 bp
* 191476 201473: contig of 9998 bp in length
* 201474 201573: gap of 100 bp
* 201574 202307: contig of 734 bp in length
* 202308 204878: gap of 100 bp
* 204879 204978: gap of 100 bp
* 204979 213531: contig of 8553 bp in length
* 213532 213631: gap of 100 bp
* 213632 218109: contig of 4478 bp in length
* 218110 218209: gap of 100 bp
* 218210 219800: contig of 1591 bp in length
* 219801 219900: gap of 100 bp
* 219901 220965: contig of 1065 bp in length.

FEATURES
Location/Qualifiers
1..220965
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p21.3"
/clone="RP6-32923"
BASE COUNT 62577 a 41674 c 45661 g 65453 t 2600 others
ORIGIN

Query Match 34.7%; Score 104.4; DB 2; Length 220965;
Best Local Similarity 94.7%; Pred. No. 3.4e-20;
Matches 108; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 58 ACACGACCTTCCTCGAATGATGGCTCCCTCAATCTATCTTTGACAAAATCTACAGCACC 127
|||||
DB 24786 ACTTGACTTCCTCGAATGATGGCTCCCTCAATCTATCTTTGACAAAATCTACAGCACC 24727
|||||

QY 128 AAGAGGACGCTGTGTCTAGCGATATGCTGTGGAAAATCTTCCTTAGGT 181
|||||

DB 24726 AAGAGGACGCTGTGTCTAGCGATATGCTGTGGAAAATCTTCCTTAGGT 24673
|||||

RESULTS
AC104439/c
LOCUS AC104439 197275 bp DNA linear PRI 20-JUN-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-793E15, complete sequence.
ACCESSION AC104439 AC024739
VERSION AC104439.2 GI:21490240
KEYWORDS H.G.
SOURCE HUMAN.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 197275)
AUTHORS Kawi, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saeedimachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
et al; Haugen, E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197279)
AUTHORS Kawi, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.

TITLE
JOURNAL

Direct Submission
Submitted (11-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 197279)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.

TITLE
JOURNAL

Direct Submission
Submitted (20-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jun 20, 2002 this sequence version replaced gi:17486621.

COMMENT

Genome Center
Center: University of Washington Genome Center

Center Code: UWGC

Web site: <http://www.genome.washington.edu>

Contact: uwgts@u.washington.edu

Drafting Center: WOGSC

Project Information

Center project name: chr-3

Center clone name: Rp11-793E15 (bc0564)

Summary Statistics

Sequencing vector: unknown; 52% of reads

Sequencing vector: plasmid; L08752; 48% of reads

Chemistry: Dye-terminator ET; 94% of reads

Chemistry: Dye-terminator Big Dye; 6% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 197168 bases at least Q40

Consensus quality: 197255 bases at least Q30

Consensus quality: 197275 bases at least Q20

Insert size: 197279; sum-of-contigs

Quality coverage: 8.2x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': Rp11-91E8 (UWGC:bc0216) AC026349

3': CTD-2563A18 (UWGC:bc0730)

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:

all regions were either double-stranded or sequenced with an

alternate chemistry or covered by high quality data (i.e., Phred

quality >= 30); an attempt was made to resolve all sequencing

problems, such as compressions and repeats; all regions were

covered by at least one plasmid subclone or more than one M13

subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest

fingerprinting. Comparison of the experimentally derived digest

fragments with sequence-predicted fragments is given below.

The electronically-digested sequence consists of both insert and

vector, in order to accurately represent the entire circular BAC.

Small fragments below a variable cutoff (approximately 400-600 bp)

are not resolved in the fingerprint and hence do not appear

in the table. There are no significant remaining discrepancies

between the experimental and predicted values. Uniquely ordered

fragments are separated by dashed lines.

HindIII

BglII

EcoRI

SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

2687 2617 8949 8586 8696 8661

6382	6410	2067	2160	6	<800
512	<800	7846	7940	2742	2803
449	<800	3734	3895	5376	5324
2602	2763	1334	1301	1493	1478
2590	2617	2287	2309	823	835
8313	8291	1814	1918	1962	2002
1711	1683	691	<800	2900	2954
9621	9472	5477	5348	1484	1478
516	<800	305	<800	1005	995
8587	8291	25245	25541	1181	1171
7446	7581	3988	4121	18560	19002
2088	2075	1633	1598	3603	3579
2509	2617	631	<800	4943	5076
3519	3501	90	<800	3239	3241
26	<800	402	<800	953	995
925	933	3350	3490	1621	1615
98	<800	4577	4515	8627	6900
6409	6410	1229	1183	79	<800
1383	1376	4221	4331	16418	16263
16048	15829	3980	4121	3361	3241
1876	1877	2548	2497	872	881
4169	4068	674	<800	4220	4193
1674	1683	2332	2309	2648	2803
52	<800	11445	11045	2283	2318
5227	5082	5692	5632	3211	3241
15464	15829	1385	1414	4615	4632
14333	14296	608	<800	6823	6900
1026	1026	3930	3895	886	881
645	<800	1465	1414	2032	2002
6621	6410	2382	2497	8680	8661
8413	8291	747	<800	5163	5076
1401	1376	1951	2012	1547	1478
1512	1471	642	<800	7058	6900
5801	5644	21060	21003	1647	1615
3256	3278	3700	3666	2753	2803
3448	3501	1321	1301	162	<800
2738	2763	10705	10502	1078	1093

LOCUS MMEWRK2 5482 bp mRNA linear ROD 07-JUL-1994
 DEFINITION M. musculus EmrK2 cDNA.
 ACCESSION X78568
 VERSION X78568.1 GI:510664
 KEYWORDS EmrK2 gene; receptor kinase.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Chai, K., Wall, C., Hanratty, R. and Keller, G.
 1 (bases 1 to 5482)
 Isolation of a gene encoding a novel receptor tyrosine kinase from differentiated embryonic stem cells
 Oncogene 9 (4), 1261-1266 (1994)
 JOURNAL
 MEDLINE
 PUBMED 94181281
 8134130
 2 (bases 1 to 5482)
 Chai, K.
 Direct Submission
 TITLE Submitted (05-JUL-1994) Chai K., 1400 Jackson St. #501, Denver Co 80206, USA
 JOURNAL
 FEATURES
 Location/Qualifiers
 1..5482
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_type="embryonic stem cells"
 249..4241
 /gene="EmrK2"
 249..4241
 /gene="EmrK2"
 /codon_start=1
 /product="embryonic receptor kinase"
 /protein_id="CAA55311.1"
 /db_xref="GI:510665"
 /db_xref="MGD:MG1:95558"
 /db_xref="SWISS-PROT:P35569"
 /translation="MVSCWQAVLPVALLGCLLIGVSGSKLVKVPESLKGIVQWQ
 AQGLFLKCRGEAAHNSLPTVVSQEDKSLTIIPSCGGRNQCSTLILTAQNH
 TGLYCRPLPTSKKKAESSYIPVSDAGSPFHEWIDIPKLVMTGROLITCRV
 TPNVTITLKEPDITPDQRIWDRRGFIANATYKEIGLNCFAVNGRLVQIN
 YLTHQTITLDVQIRPSPVRLILGQILVNCATTELNTRVQSMNYPGRATKRS
 IKQIDRSHNNVHRSVLKINNVSRDGLYICRVKSGSFOSNTSHVYKAFIS
 VKRKPQVETTAGRSYRLSKMKVAFSPETVLKQSPATKASRYLHVYSLLIK
 DVTEDAGYITLIGIKQSLFNALITLVNKPQYKASVLSPLPLPGSKQV
 LUCYVGPRTITLWHPCHNHSKERYDFCTENEEFILDPSLNGRIEISQR
 VIEGNTVSTLVADSPQGIYSCRAFNKIGIVERNIKFYVDVPGFHVSLKMP
 AEGEDKLSCVYNKFLYRDTITLWLTNNRTHMNSTSKQNHSSYSTILNLVKN
 SLEDSGTACRANITYTGEDILAKTEVLVROSEAPHLIQLNSDVEYSISGTLICQA
 RVPAPQITLKNKHKIQBPGLILQPGNSTLFIERTVEDEGVRCRSTNKGAVES
 AAYLVQGTSDSKNLELITLTCTVAAITLFIITLITRKLKSSSEVTDYLSIM
 PDEVLDEQCELYDASNEFARERKLKSLGRGAFKGVQASAFKKSPTCRV
 AVKLKKEATASEYKALMTLTKLTHIGHLVNVLGACTKOGPLWIVYCYKGN
 LSNYLKSRDLQKDAALHMLKESLEPGLEGQKPRDVSSTSSFTFEDR
 SVSDVEGEDYSEISKQPLTMEDLISYEQVARGFELSKKCIHRDLAARNILLSEN
 NVYKICDGLARDIYKNPVYVRGDTIRLPKMAPESTFOKYSIKSDVMSYGVILWE
 IFSLGSGYPQVQMDQFCSLKEGMRMTPEYATPEIYQIMLDQWKPKEPFEAE
 LVEKLGLDLOANVQDQKDIPLNAILTKNSFTYPTISEDRGCGADPHEHSGS
 SDQRYVYNAKFMASLERIKTFEELSPNSTSMEDYGLDITSLIGSLPKRFWITKP
 SKMLIDIRIAPSKKEAGLSDLPRPSPFCSSCGHIRPVQDDESELKSCCSPPEYDN
 SVLYSSPPA"
 BASE COUNT 1464 a 1446 c 1369 g 1203 t
 ORIGIN
 Query Match 31.8%; Score 95.8; DB 10; Length 5482;
 Best Local Similarity 89.6%; Pred. No. 1.2e-17;
 Matches 103; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 66 AGACATCGACTTCCTCTGAATGGATGGCTCTGATCTATCTTGACAAATCTACAGCA 125
 ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
 Db 3412 ATATCTGACTTCCTCTGAATGGATGGCTCTGATCTATCTTGACAGGCTACAGCA 3471
 ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
 QY 126 CCAAGAGCGAGTGTGCTCTACGAGGATATGCTGTCGGGAAATCTCTCCTTAGG 180
 ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++

Db 3472 CCAAGAGCGAGTGTGCTCTACGAGGATATGCTGTCGGGAAATCTCTCCTTAGG 3526
 ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
 RESULT 11
 MUSRIKB
 LOCUS
 DEFINITION Mus musculus receptor tyrosine kinase (FLI) mRNA, complete cds.
 ACCESSION L07297
 VERSION L07297.1 GI:293782
 KEYWORDS receptor protein tyrosine kinase.
 SOURCE Mus musculus (strain C57BL/6J, sub-species domesticus) cDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 6055)
 Flitzner, H., Kelleher, K., Morris, G.E., Bean, K., Marberg, D.M., Kitz, R., Morris, J.C., Sookdeo, H., Turner, K.J. and Wood, C.R.
 Molecular cloning of murine FLT and FLT4
 Oncogene 8 (8), 2293-2298 (1993)
 5330572
 8330572
 Location/Qualifiers
 1..6055
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /sub_species="domesticus"
 /db_xref="taxon:10090"
 1..6055
 /gene="FLT"
 27..4028
 /gene="FLT"
 /note="putative"
 /codon_start=1
 /product="receptor tyrosine kinase"
 /protein_id="AAA40078.1"
 /db_xref="GI:293783"
 /translation="MVSCWQAVLPVALLGCLLIGVSGSKLVKVPESLKGIVQWQ
 AQGLFLKCRGEAAHNSLPTVVSQEDKSLTIIPSCGGRNQCSTLILTAQNH
 TGLYCRPLPTSKKKAESSYIPVSDAGSPFHEWIDIPKLVMTGROLITCRV
 TPNVTITLKEPDITPDQRIWDRRGFIANATYKEIGLNCFAVNGRLVQIN
 YLTHQTITLDVQIRPSPVRLILGQILVNCATTELNTRVQSMNYPGRATKRS
 IKQIDRSHNNVHRSVLKINNVSRDGLYICRVKSGSFOSNTSHVYKAFIS
 ASVIRKPQVETTAGRSYRLSKMKVAFSPETVLKQSPATKASRYLHVYSLLIK
 DVTEDAGYITLIGIKQSLFNALITLVNKPQYKASVLSPLPLPGSKQV
 LUCYVGPRTITLWHPCHNHSKERYDFCTENEEFILDPSLNGRIEISQR
 VIEGNTVSTLVADSPQGIYSCRAFNKIGIVERNIKFYVDVPGFHVSLKMP
 AEGEDKLSCVYNKFLYRDTITLWLTNNRTHMNSTSKQNHSSYSTILNLVKN
 SLEDSGTACRANITYTGEDILAKTEVLVROSEAPHLIQLNSDVEYSISGTLICQA
 RVPAPQITLKNKHKIQBPGLILQPGNSTLFIERTVEDEGVRCRSTNKGAVES
 AAYLVQGTSDSKNLELITLTCTVAAITLFIITLITRKLKSSSEVTDYLSIM
 PDEVLDEQCELYDASNEFARERKLKSLGRGAFKGVQASAFKKSPTCRV
 AVKLKKEATASEYKALMTLTKLTHIGHLVNVLGACTKOGPLWIVYCYKGN
 LSNYLKSRDLQKDAALHMLKESLEPGLEGQKPRDVSSTSSFTFEDR
 SVSDVEGEDYSEISKQPLTMEDLISYEQVARGFELSKKCIHRDLAARNILLSEN
 NVYKICDGLARDIYKNPVYVRGDTIRLPKMAPESTFOKYSIKSDVMSYGVILWE
 IFSLGSGYPQVQMDQFCSLKEGMRMTPEYATPEIYQIMLDQWKPKEPFEAE
 LVEKLGLDLOANVQDQKDIPLNAILTKNSFTYPTISEDRGCGADPHEHSGS
 SDQRYVYNAKFMASLERIKTFEELSPNSTSMEDYGLDITSLIGSLPKRFWITKP
 SKMLIDIRIAPSKKEAGLSDLPRPSPFCSSCGHIRPVQDDESELKSCCSPPEYDN
 SVLYSSPPA"
 BASE COUNT 1653 a 1549 c 1492 g 1381 t
 ORIGIN
 Query Match 31.8%; Score 95.8; DB 10; Length 6055;
 Best Local Similarity 89.6%; Pred. No. 1.2e-17;
 Matches 103; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 66 AGACATCGACTTCCTCTGAATGGATGGCTCTGATCTATCTTGACAAATCTACAGCA 125
 ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
 Db 3199 ATATCTGACTTCCTCTGAATGGATGGCTCTGATCTATCTTGACAGGCTACAGCA 3258
 ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++

Tel:81-3-5449-5550, Fax:81-3-5449-5425)

FEATURES

source

Location/Qualifiers

1. 4272

/organism="Gallus gallus"

/db_xref="taxon:9031"

/tissue_lib="whole embryo"

1. 4272

/gene="vegfr-1"

267. 4250

/gene="vegfr-1"

/note="flt-1"

fms-like tyrosine kinase-1"

/codon_start=1

/product="vascular endothelial growth factor receptor-1"

/protein_id="BAB84690.1"

/db_xref="GI:18652861"

/translation="MPROLLSGTVLLGAAPFLLAGSTSGSKLVPVLSVNGRQHWVAG

QTLNLCRGEMLSHSLPEALSKDSKRLNVTYACGRNGICQFCSTLTLSTLSCANDICR

YSCRYPTSPVKKKRESIVVFINDISNPFVEHSDIPKLIHWTYKEMILPCRYVAPN

IAYTLAKIPREVLIDGGYTIIDWNRGFRIPENYAFIIGLSCETIIGGRKISIKVIT

HRENTIFDKLSTPRVLKLLGDSLAINCTVAAANRVQMTWTFEGEAKKRSVIQ

RIDCKREANVYSILVDKVRDIDKQVACHKSPKSNKLVITVIVYKRFINLKR

RDKTLEAVAGSKYRLEPMKYKAFSPVETWLDGLPAEKACARYMKNYSLLIKDVA

EDAGNYTILSLRQWNLKNIIVTLKVNKPOIVENAVSFPDPNLYLLSSKQVILIC

TVYGIPTTIVWVPCRNHSTKPRGFCSTGSDVFNKLTGNTGNISNLIQSIIEIRI

EGKNVASTLVAAEAKSSGIYSCVANSKYKAPRVNSELVDPVSGPHISLEKVPLEG

ENLVLSCKANPKMYKDIJSHLIPREVINOZKAKKALNKEYSITLITIRNLSLSRGYI

TCRARNIFGKVLQKQVSTRAGEAPALLQLMDQTVNTSNALECCQVHGIFEPQI

TWKNHEEIQQSGHILGPGSRMLFIERVKEDEGLYCIATNLKGSVESTAYIVQGL

TVRSNLELITCTCVAATFLWLLFLFKLRPFSETKTNHLSILIMLDEIVPL

DEQECIPYDASKIARERLKGSGHGFAGKVVQASAFGIKKPTCRIVAVKMLK

EGATASEYKALMTELKILIHIGHNLNVLGACKNGGPIWVVEYCKVGNLSNVLK

SKRNFSPKDSLOGLMDKKIGIEPVKKORLASVTSSEFASSGFEDSLSDA

FEDEDAELEYKLPTMEDLLISVFGVARGHEFLSRKCIHRDLAANLLSNVYK

ICDFGLARIYKNPDYVRGDARLPLKMAPEISFDKTYNTKSDVSYGVLLSEIFSR

GAPYFGVQIDDFCSKLKTEGRNRAPEATEEIVYIMLDCRNSPNRERWSELYK

LGDLQAQQQGGDYIPLDTIFAESGPPASDPLCNEKFPVSPNCRSTERARIN

TFKIKPPQRIKFEELPIKEKLVNDYQADSGWLASBELKRLTWGSKOKWLPQMK

GVSRKESGLSITKPRSFCSQDLSKSKRYTGNVLEKMKACHSPPPPSVSV

HYSOPI"

BASE COUNT 1266 a 1001 c 1083 g 922 t

Query Match 29.2%; Score 88; DB 5; Length 4272;

Best Local Similarity 86.6%; Pred. No. 2.5e-15;

Matches 97; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 69 CTCGACTCCCTGAAATGATGGCTCTGAAATCTATCTTTGACAAATCTACACACCA 128

3421 CTCGACTCCCTCAAGTGGATGGCTCCAGAAATCCATCTTTGTATAAATCTACACACCA 3450

Qy 129 AGAGCGAGCTGTGCTTACGAGTATTGCTGTGGAAATCTCTCTTAGG 180

Db 3481 AAAGCGATGTGGTCTACGGGTCTTCTCTGGGACATATTTCTCTGGG 3532

RESULT 15

AL646088

LOCUS

DEFINITION

AL646088

ACCESSION

AL646088.23

VERSION

AL646088.23

KEYWORDS

HTG.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (21-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

COMMENT

humquers@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 25, 2002 this sequence version replaced gi:20520452.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP
database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-58E13 is
from the RP21-23 Mouse PAC Library
constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6

Location/Qualifiers

1. 211330

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="11"

/clone="RP23-58E13"

/clone_lib="RP21-23"

BASE COUNT 63824 a 45513 c 46613 g 55380 t

ORIGIN

Query Match 26.8%; Score 80.6; DB 10; Length 211330;

Best Local Similarity 67.7%; Pred. No. 4.1e-13;

Matches 113; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 50 CCCCTTGCCCTCTCTAGACTCGACTTCCTCTGAAATGGATGGCTCTGAAATCTATCTTT 109

Db 189669 CTCCTTTTGGCCACCCAGCCGACTTCCTCTGAAATGGATGGCTCTGAAATCTATCTTT 189728

Qy 110 GACAAAATCTACAGCAACAGAGCGAGTGTCTTACGAGTATTGCTGTGGAAATC 169

Db 189729 GATAGGTGTACACACGACAGATGATGCTGTGGCTGTCTGTGGGAGATC 189788

Qy 170 TTCTCCTTAGTAAATTTGGAGAGGAGAAATCAACAGCCACCA 216

Db 189789 TTCTCATTTGGTGTGATGAGAGTGGTGGCCATCCAGGAGAGAAGA 189835

Search completed: December 8, 2002, 07:02:25

JOB time : 2754.03 SECS

GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

QM Nucleic - nucleic search, using sw model

Run on: December 8, 2002, 11:59:28 ; Search time 117.413 Seconds

(without alignments)

6977.502 Million cell updates/sec

Title: x51602_COPY_1900_4000

Perfect score: 2101

Sequence: 1 TATATCACATGTCGCAAA.....CGACACACACATGTCGTCG 2101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Residues: 350425 seqs, 194966369 residues

All number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCI_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	710	33.8	5470	10	US-09-766-678-1
2	708.4	33.7	5406	10	US-09-919-408-5
3	708.4	33.7	5406	10	US-09-872-136-5
4	541.6	25.8	4425	10	US-09-982-610-31
5	541.6	25.8	9108	10	US-09-982-610-45
6	282.8	13.5	6827	10	US-09-982-610-17
7	259	12.3	5084	9	US-09-954-531-1383
8	259	12.3	5084	10	US-09-967-768A-277
9	249.4	11.9	3501	10	US-09-919-408-3
10	249.4	11.9	3501	10	US-09-872-136-3
11	247.8	11.8	3120	10	US-09-982-610-22
12	233.6	11.1	3992	10	US-09-944-807-9
13	221.6	10.5	3453	10	US-09-919-408-1
14	221.6	10.5	3453	10	US-09-872-136-1
15	213.4	10.2	3270	10	US-09-866-510-1
16	213.4	10.2	3270	10	US-09-866-510-9
17	213.4	10.2	4054	9	US-09-955-363-35
18	213.4	10.2	6378	10	US-09-919-497-40
19	213.4	10.2	6412	10	US-09-769-987-1

20	211.8	10.1	3270	10	US-09-866-510-3
21	211.8	10.1	3270	10	US-09-866-510-7
22	210.2	10.0	3270	10	US-09-866-510-5
23	194.4	9.3	3454	12	US-10-044-050-48
24	194.4	9.3	3726	10	US-09-925-302-271
25	184.6	8.8	4268	10	US-09-954-456-293
26	184.6	8.8	4268	10	US-09-954-456-1595
27	175.8	8.4	3321	10	US-09-866-510-21
28	175.8	8.4	3321	10	US-09-866-510-13
29	177.2	8.4	3321	10	US-09-866-510-15
30	177.2	8.4	3321	10	US-09-866-510-17
31	177.2	8.4	4465	9	US-09-955-363-1
32	177.2	8.4	4465	10	US-09-955-363-1
33	175.6	8.4	3321	10	US-09-728-952-82
34	166.6	7.9	392	10	US-09-960-352-4233
35	166.6	7.9	402	10	US-09-960-352-1442
36	166.6	7.9	391	10	US-09-960-352-2603
37	148.2	7.1	2782	9	US-10-108-605-250
38	132	6.3	415	10	US-09-960-352-10423
39	128.6	6.1	366	10	US-09-960-352-10878
40	127.8	6.1	434	10	US-09-960-352-5227
41	126.4	6.0	2869	9	US-10-016-283-2
42	126.8	5.7	410	10	US-09-960-352-4303
43	119	5.7	338	10	US-09-962-436-555
44	118.4	5.6	2431	12	US-10-036-869-35
45	111	5.3	3227	10	US-09-919-497-2

ALIGNMENTS

RESULT 1

US-09-766-678-1
 : Sequence 1, Application US/09766678
 : Patent No. US20020081650A1
 : GENERAL INFORMATION:
 : APPLICANT: Ullrich, Axel
 : Millauer, Birgit
 : Gazit, Aviv
 : Levitzki, Alex
 : TITLE OF INVENTION: Flk-1 Is A Receptor For Vascular
 : Endothelial Growth Factor
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennle & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : Zip: 10036-2711
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/766,678
 : FILING DATE: 25-Jan-2001
 : CLASSIFICATION: <UNKNOWN>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/193,829
 : FILING DATE: 09-FEB-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Coruzzi, Laura A.
 : REGISTRATION NUMBER: 30,742
 : REFERENCE/DOCKET NUMBER: 7683-060
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (212)750-9050
 : TELEFAX: (212)869-9741
 : TELEX: 66141 PENNIE
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:

SEQUENCE 3, Appli
 Sequence 7, Appli
 Sequence 5, Appli
 Sequence 48, Appli
 Sequence 271, App
 Sequence 293, App
 Sequence 1599, Ap
 Sequence 21, Appli
 Sequence 13, Appli
 Sequence 15, Appli
 Sequence 17, Appli
 Sequence 1, Appli
 Sequence 82, Appli
 Sequence 19, Appli
 Sequence 4293, Ap
 Sequence 14442, A
 Sequence 2603, Ap
 Sequence 250, App
 Sequence 10423, A
 Sequence 10878, A
 Sequence 5227, Ap
 Sequence 2, Appli
 Sequence 4303, Ap
 Sequence 555, App
 Sequence 35, Appli
 Sequence 2, Appli

```

; LENGTH: 5470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 286..4386
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-766-678-1

Query Match 33.8%; Score 710; DB 10; Length 5470;
Fast Local Similarity 65.0%; Pred. No. 5.9e-200;
Matches 1081; Conservative 0; Mismatches 575; Indels 2;

QY 201 CACTCTATCTTACATCATGATGTTCCCTGCAAGATTCAGGACCATATGCCTGGCAG 260
DB 2148 CATCTTGATGTGGCATTTCAAGATGCCCTCTGAGGACCAAGGCGACTATGTTTGTCT 2207
QY 261 AGCCAGGAATGTATACACAGGGGAAGAAATCCCTCCAGAGAAAGAAATACAAATCAGAGA 320
DB 2208 TGCTCAAGTANGAGACCCAGAAAGACATGCCCTGTCAACAGCTCATATCTAAG 2267
QY 321 TCAGGAGACCATACCTCCCTGCGAAACCTCAGCTGATCAGACAGTGGCCATCAGCAGTTC 380
DB 2268 GCGCATGGACCATGATCACCGGAATCTTGGAGAAATCAGACAAACCAATGGCCAGAG 2327
QY 381 CACCACCTTAGCTGCTGATGCTATGATGTCGCCGAGCTCAGATCACTTGGTTTAAAGA 440
DB 2328 CATTCAGTACCTGCCAGCATCTGGAATCTTACCCACACATTCATGGTTTCAAGA 2387
QY 441 CAACCAAAAATACAAAGAGACCTCGGAATATTTTAGGACCAAGAGCAGCAGCGTGT 500
DB 2388 CAACGAGACCTGGTGAAGATTCAGGCATTTGCTAGAGATGGGAACCGGAACCTGAG 2447
QY 501 TATGAAAGAGTCACAGAGAGATGAGTGTCTATCACTTGCACAGAGAGAGAGAGAGAGAG 560
DB 2448 TATCCGAGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2507
QY 561 GGGCTCTGTGAAAGTTCAGATACCTCACTGTTCAGGAACCTCGCAGCAAGTCTAATCT 620
DB 2508 TGCTGTGACAGAGCGGAGAGCGCTCTTCATATAGAGAGTCCCGAGAAAGACCACT 2567
QY 621 GGAGCTGATCACTTAAACAGCAGCTGTGTGGCTCGGACTCTCTCTGGCTCTATTAAC 680
DB 2568 GGAAGTCATATCTCTCGTGGCACTGAGTGTGCTATGCTTCTTCTGGCTCTCTCT 2627
QY 681 CTCTCTTATCCGAAATGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 737
DB 2628 CATGTCTTACGGACCGTTAGCGGGGCGCAATGAAGGGGAGCTGAAGACGGCTACTTGT 2687
QY 738 AATTATTAATGGACCATGATGATGCTTGTGATGAGCAGTGTGAGCGGCTCCCTTATGA 797
DB 2688 TATTGTTCATGGATCCAGATGAATGGCCCTTGGATGAGCGCTGTCAACGCTTGGCTTATGA 2747
QY 798 TCCGACAGATGGGAGTTTGGCCGGAGAGACTTAAACTGGGCAATCACTTGGAGAGAG 857
DB 2748 TGGCAGCAATGGGAAATTCGCCAGGCGGTGAAGACTGAAAGAAACCTCTTGGCCGGAG 2807
QY 858 GGGTTTGGAAAGTGGTTCAAGCATCAGCATTTGGCATTGAAGAAATCACTTACGTGGCG 917
DB 2808 TGCTTCGGCCAGTATTGAGGAGAGAGCTTTTGGATTCAGAGACACGCGACTTGGCAA 2867
QY 918 GACTTGGCTGTGAAGTCTGTAAGAGAGGGGCCAGCGCCAGCGAGTACAAAGCTGTGAT 977
DB 2868 AACAGTAGCGGTCAAGATGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2927
QY 978 GACTGACCTTAAATCTTACCCACATTCGCCACCATCTGAACGTGTTAACTGCTGGG 1037
DB 2928 GTCTGAACTCAAGATCTCTCATCCATTTGTCACCACTTCAATGTGTTGACCTCTAGG 2987
QY 1038 AGCCTGCACCAAGCAGGAGGCGCTCTGATGTGATGTTGTAATCTGCAAAATATGAAA 1097

```

```

DB 2988 CGCTCGACCAAGCCGGAGGGGCTCTCACTGATTTGGAAATCTGCAAGTTTGGAAA 3047
QY 1098 TCTCTCAACTACTCTCAAGAGCAAAAGTACTTATTTTCTCACAAGATGCACTACT 1157
DB 3048 CTTATCACTTACTTACGGGCAAGAGAAATGAATTTGTTCCCTATAGAGCAAGGGGC 3107
QY 1158 ACACATGGAGCCTAAGAGAGAAATAGGAGCCAGGCCCTGGAAACAAAGCAAGCAAG 1217
DB 3108 AGCTTCCCGCAGGCAAGGACTACGTGGG---GAGCTCTCCGAGATCTGAAAGAGG 3164
QY 1218 ACTAGATAGCTCACCAGCAGCGAAAGCTTTGCGAGCTCCGCGCTTTCAGGAAGATAAG 1277
DB 3165 CTGGACAGATCACCAGCAGCAGGCTCTGCCAGCTCAGGCTTTCTTGGAGAGAAATC 3224
QY 1278 TCTAGTGTGTTGAGGAGAGAGGATTTCTGACGTTTCTACAGAGAGCCCATCATAT 1337
DB 3225 GCTCAGTGTGTTGAGGAGAGAGAGGCTTCTGAAGAACTGTACAGGACTTCTGACCTT 3284
QY 1338 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGGCAATGGATTCCTGTTCCAG 1397
DB 3285 GGAGCACTCATCTGTTACAGCTTCCAGTGGCTTAGGCAAGAGGATCTTGCAATCAG 3344
QY 1398 AAGTGCATTCATCGGACCTTGGCAGCGAGAAACATTTCTTATCTGAGAAACACCTGT 1457
DB 3345 GAAGTGTATCCACAGGAGCTTGGCAGCAACCAACATTTCTTATCGGAGAGAAATGT 3404
QY 1458 GAAGATTTGATTTTGGCTTGGCCGGGATTTATTAAGAACCCCATATATGTGAGAAA 1517
DB 3405 TAGATCTGTGACTTGGCTTGGCCGGGACATTTATTAAGACCCGGAATATGTGAGAAA 3464
QY 1518 AGGAGATCTGACTTCTCTGAAATGGATGGCTCCCGAAATCTATCTTTCACAAAATCTA 1577
DB 3465 AGGAGATGCGCGACTCCCTTTGAGTGGATGGCCCGGAGAACATTTTTCACAGATTA 3524
QY 1578 CAGCAGAGAGAGAGTGTGGTCTTACGAGTATGCTGTGGGAAATCTTCTCTTAGG 1637
DB 3525 CACAATTCAGACGATGTGTGCTTTCGTTGCTGCTGCGGAAATTTTCTCTTAGG 3584
QY 1638 TGGCTTCCATACCCAGGAGTACAAATGGATGGAGCTTTTGAATCGCTGAGGAGAGG 1697
DB 3585 TGCTCCCGTACCTTGGGTCAGATGATGAAGAAATTTTGTAGAGATTTGAAGAAGG 3644
QY 1598 CATGAGATGAGAGCTCTGAGTACTTACTCTGAAATCTATCAGATCATGCTGACTGT 1757
DB 3645 AACTAGATGCGGCTCTGCTGACTACTTACCCAGAAATGTTACAGACCATGCTGACTGT 3704
QY 1758 CTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1817
DB 3705 CTGGCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3764
QY 1818 TTTGCTTCAAGCAATGTACAAAGAGATGTTAAAGACTACAT 1859
DB 3765 CTTCTGCAAGCAATGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3806

```

```

RESULT 2
US-09-766-678-1
; Sequence 5, Application US/05919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Inclone Systems Incorporated
; STREET: 180 Varlick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: JN 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/674,666
FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM 3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRAINEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-919-408-5

Query Match 33.78; Score 708.4; DB 10; Length 5406;
Best Local Similarity 55.04; Pred. No. 1.7e-199;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;

QY	201	CACCTCTTAATCTACCATCAATGATGTTCCCTGCGAAGATTCAGGCACCTATGCTTCGAG	260
DB	2070	CATCTTAATGTCGCAATTCAGATGCTCTGCGAGACCAAGCGCACTATGTTGCTC	2129
QY	261	AGCCAGCAATGTATACAGAGGGAAGAAATCTCCAGAGAGAGAAATTAACAATCAGAGA	320
DB	2130	TGCTCAGATAGAGACCAAGAGAGAGATGCTTGGTCAACAGCTCATCATCTCCTAGA	2189
QY	321	TCAGAGAGCAATACCTCTCGGAAACCTCAGTATCAGACAGTGGCCATCAGCAGTTC	380
DB	2190	CGCATGCGCAATGATCAGCGGAATCTCGAGATCAGACAAACCACTTGGCGAGC	2249
QY	381	CACCACCTTTAGCTGTATGCTATGCTGTCGCCAGGCTCAGATCCTTGTTTAAAAA	440

DB	2250	CAITGAAGTGAATGCTGCTGAGCATCTGGAATCTACCCACACATTAATGTTTCAAGA	2309
QY	441	CAACCAACAAATACAAACAGAGCCTGTGAATATTTTAGGACAGGAGCAACAGCGTGT	500
DB	2310	CAAGAGACCTCGTGTAGATGATGAGCATTTGCTGAGAGATGGAACCGAACCTGAC	2369
QY	501	TATGAAGAGTGCACAGAGAGAGATGAAGTGTCTATCACTGCAAAAGCCCAACACAGAA	560
DB	2370	TATCCGACGGTGTAGAGAGAGATGAGAGCCTCTACACCTGCCAGGCTTCAATGTCT	2429
QY	561	GGGCTCTGTGGAAGTTCAGCATACCTCACTGTTCAGGAACCTCGGACAGCTATATCT	620
DB	2430	IGGCTGTGCAAGAGCGGAGACGCTCTCATANTAGAGGTGCCAGGAAAGACCAAT	2489
QY	621	GGAGCTGATCACTCTAAACATCACCCTGTGTGGTGGCAGCTCTCTTCTGGCTCTTAAAC	680
DB	2490	GGAACTCATATCTCTGCTGGCAGCTGAGTGGCATGTTCTTCTGGCTCTTCTGT	2549
QY	681	CCCTCTTAICCGAAATGAAAGGTC---TCTTCTGAATAAAGACTGACTACCTATC	737
DB	2550	CATCTGTGTACGACGCTTAAGCGGGCAATGAAGGGGAACCTGAAGACAGCTACTITG	2609
QY	738	AATTAATAGGACCCAGATGAAGTTCCTTTGGATGAGCAGTGTGAGCGCTCCCTTAIGA	797
DB	2610	TATGTCAIGATCCAGATGATGCTTGGATGAGCGCTGTGACGCTTGCCTTAIGA	2669
QY	798	TGCCAGCAAGTGGGAGTTTGGCCGGGAGACACTTAACTGGGCAATCCTTTGGAGAGS	857
DB	2670	TGCCAGCAAGTGGGAAATCCCGAGGACCGCTGAAACTAGGAAACCTCTTGGCGCGG	2729
QY	858	GGCTTTTGAAAGTGGTTCAGCATCAGCAITGGCATTAGCAATCACCCTAGCTGCCG	917
DB	2730	TGCTTCGGCCAAAGTGTGAGCGAGCGCTTTTGGAAATGACAGACGCGACTTGCA	2789
QY	918	GACTGTGCTGTGAAATGCTGAAGAGAGGGGCGCCAGCGCAGAGTACAAAGCTGTAT	977
DB	2790	AACAGTAGCGTCAAGATGTTGAAGAGAGGCAACACACAGCAGCAICGAGCCCTAT	2849
QY	978	GACTGAGCTAATAATCTTGACCCACATTTGCCACCATCTGAACGTGGTTACCTGTGG	1037
DB	2850	GCTGSAACCTCAAGATCTCTCCATTTGTCACCATCTCAATGTGTGACCTCTCTAGS	2909
QY	1038	AGCTGTGACCAAGCAGGAGGCGCTCTGATGGTGTGTTGTAATGACTGCAATATGNA	1097
DB	2910	CGCTGTGACCAAGCGGAGGCGCTCTGATGGTGTGTTGTAATGACTGCAATATGNA	2969
QY	1098	TCTCTCACTACTCTCAAGAGCAACGCTGACTTATTTTCTCAACAGAGATGACGACT	1157
DB	2970	CCTATCAACTTACTTACGGGCAAGAGAAATGAATTTGTTCCCTATAGAGCAAGGGGC	3029
QY	1158	ACACATGAGCGCTTAAGAGAGAAATGAGCGGCGCTGGAACAGCAAGCAAGCAAG	1217
DB	3030	ACGCTTCCGCGAGGCAAGGACTACGTTGGG---GAGCTCTCCGTGGATCTGAAGAGCG	3086
QY	1218	ACTAGATAGCGTCAACAGCAGCAAGAAAGCTTTGGGAGCTCCGGCTTTCAGGAAGATA	1277
DB	3087	CTTGACAGCATCAACAGCAGCAGGAGCTCTGCCAGCTTGTGAGGAGAAATC	3146
QY	1278	TCAGTGTGATGTGAGAGAGAGGAGTCTGACGGTTTCTTCAAGAGGAGCCATCAT	1337
DB	3147	GCTCAGTGTGATGTGAGAGAGAGGAGTCTTGAAGAACTGTACAAAGGAGTCTGACCT	3206
QY	1338	GGAAGATCTGATTTCTACAGTTTTCAGTGGCCAGGAGCATGAGTTCTGTTCCAG	1397
DB	3207	GGACATCTCATCTGTACAGCTTCCAGTGGCTTGAAGGCAATGAGTTTCTTGGCATCAG	3266
QY	1398	AAAGTGAATTCAGGAGCTGGGAGCGAGAAACATCTTTTATCTGAGAAACAGTGT	1457
DB	3267	GAAGTGTATCCACAGGAGCTGGGAGCAGCAACAACTTCTCTATCGGAGAGATGTGT	3326
QY	1458	GAACTTTGTGATTTTGGCTTGGCCGGATTTTATAAGAACCCCGATTTATGTGAGAA	1517
DB	3327	TAGATCTGTGACTTGGCTTGGCCGGAGCAITTTATAAGACCCGGATTTATGTGAGAA	3386

QY 1518 AGGCAATACGACTTCCTCTGAAATGAGTGCCTCCCGAATCACTAATCTTTCACAAATCTA 1577
Db 3387 AGGATCCCGGCTCCCTCTGAAATGAGTGCCTCCCGAATCACTAATCTTTCACAGAGTAA 3446
QY 1578 CAGCACCAGAGGAGGCTGCTTACGAGTATTCCTGCGGAAATCTTCTTCCTTAGG 1637
Db 3447 CACAATTCAGAGGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3506
QY 1638 TGGGTCTCCATACCCAGAGTACAAATGATGAGAGCTTTTCAGTCGCTCAGGGAAGG 1697
Db 3507 TGGCTCCCATACCCAGAGTACAAATGATGAGAGCTTTTCAGTCGCTCAGGGAAGG 3566
QY 1698 CATGAGATGAGAGCTCTGAGTACTTACTCTGAAATCACTACATCACTGCTGCTG 1757
Db 3567 ACTAGATGAGAGCTCTGAGTACTTACTCTGAAATCACTACATCACTGCTGCTG 3626
QY 1758 CTGGCAGAGAGCCCAAGAAAGGAGGAGATTTTCAGAGTCTTGGAAAATCTAGGTGA 1817
Db 3627 CTGGATGAGAGCCCAAGAAAGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 3686
1818 TTGCTTCAAGCAATCTACAGAGTGGTAAAGACTATAT 1859
Db 3687 CTCTGCAAGCAATCTACAGAGTGGTAAAGACTATAT 3728

RESULT 3

US-09-872-136-5

Sequence 5, Application US/09872136

Patent No. US20020119545A1

GENERAL INFORMATION:

Applicant: Lemischka, Ihor R.

Title of Invention: TOPIPOENT HEMATOPOIETIC STEM CELL

RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

Street: 180 Varick Street

City: New York

State: New York

Country: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/872.136

FILING DATE: 01-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/208.786

FILING DATE: <Unknown>

APPLICATION NUMBER: US/09/021.324

FILING DATE: <Unknown>

APPLICATION NUMBER: US/07/977.451

FILING DATE: 1992-11-19

APPLICATION NUMBER: US 07/906.397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813.593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793.065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728.913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/579.665

FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28.601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-872-136-5

Query Match 33.7%; Score 708.4; DB 10; Length 5406;

Rest Local Similarity 65.0%; Pred. No. 1.7e-192;

Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;

QY 201 CACTCTTAATCTTACCAATCAATGATTTTCCTTCCCAAGATTCAGGCACTATGCTGCAG 260

Db 2070 CAITTTGATTTGGCAATTTCAAGATTCCTCTCTGAGGACCAAGGAGCATGTTTGCTC 2129

QY 261 ACCCAGGAAATGATACACAGGAGGAGAAATCTCCACAGAAAGAAATTAACAATCAGAA 320

Db 2130 TGCTCAAGATAAGAACCAACCAAGAAAGACATTCCTGCTGTTCAACAGCTCATCTCCTAGA 2189

QY 321 TCAGGAGCAGCAATCTCTCTCGAAGCTTCAGTATCACACAGTGGCCATCAGCAGTTC 380

Db 2190 GGGCATGCGACCCATGATCATCCGGAATCTCGAGAAATCAGCAACCAATTCGGAGAC 2249

QY 381 CACCACTTACATGCTATGCTTAATGTTGCTCCCGAGCCCTCAGATCACTTGGTTTAAAAA 440

Db 2250 CATGGAAGTGAATGCTGCTGAGAAATCTCTACCCACACATTAATGATGTTTCAAGA 2309

QY 441 CAACCAAAAATACAAAGAGCTGGAAATTAATTTAGGACAGGAGGAGCAGCCTGTT 500

Db 2310 CAACGAGACCTTGTAGAGATTCAGGCAATGTACTAGAGATGGGAACCGAACCTTGAC 2369

QY 501 TATTGAAGAGTCAACAGAGAGATGAAGGTGCTATCTACTGCAAGGCCACCAACAGAA 560

Db 2370 TATCCGAGGCTGAGGAAGGAGGAGTGGAGGCTCTACACCTGCCAGGCTGCAATGCTT 2429

QY 561 GGGCTCTGGAAGTTCAGCATACCTCCTCTCAGGAACTCGGACAGTCTAATCT 620

Db 2430 TGGCTGTCAAGAGCGGAGAGCTCTTCTATAGAGGTGGCCGAGGAGGAGGAGGAGGAG 2489

QY 621 GGAGCTGATCACTTAACATGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680

Db 2490 GGAAGTCAATCTCTGCTGAGCACTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2549

QY 681 CTTCTTATCGGAATGAAGGTC---TCTCTGAAATTAAGACTGACTACCTATC 737

Db 2550 CATCTCTAGGAGCCGTTAAGCGGCCCAATGAAGGAGGAGGAGGAGGAGGAGGAGGAG 2609

QY 738 AATTATGAGCCAGATGAAGTTCCTTGTGATGAGCAGTGTGAGCGGCTCCCTTATGA 797

Db 2610 TATGTCAAGTCAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2669

798 TCCAGCAAGTGGAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 857
2670 TCCAGCAAGTGGAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 2729
858 GCGTTCCTGGAAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 917
2730 TCCAGCAAGTGGAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 2789
918 GACTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 977
2790 AACAGTACGCTGGAAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 2849
978 GACTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 1037
2850 GTCTGACTGAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 2909
1038 AGCTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 1097
2910 CCGCTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGACTTAACTGGGCAAACTACTTGGACAGG 2969
1098 TCTCTCACTACTGAG 1157
2970 CCACTCACTACTGAG 3029
1158 ACATGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
3030 AGCTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3085
1218 ACTAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1277
3087 CTTGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3146
1278 TCTGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1337
3147 CTTGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3206
1338 GGAAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1397
3207 GGAAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3266
1398 AAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
3267 GAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3325
1458 GAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1517
3327 TAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3386
1518 AGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1577
3387 AGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3445
1578 CAGCAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1637
3447 CAAATGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3506
1638 TGGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1697
3507 TGGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3566
1698 CATGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1757
3567 AACTGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3625
1758 CTGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1817
3627 CTGAGTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3686
1818 TTGCTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1859
3687 CTTCTGTGCTGTGGAAGTGTTCCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

RESULT 4
US-09-982-610-31
Sequence 31, Application US/09982610
Patent No. US20020148420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPalin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222,616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 4425 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-982-610-31
Query Match 25.8% Score 541.6; DB 10; Length 4425;
Best Local Similarity 58.7% Pred. No. 4,6e-150;
Matches 981; Conservative 0; Mismatches 679; Indels 12; Gaps 2;
QY 185 CTAGGAGACCTCCATCACTCTTAATCTTACCATCAATGATGTTTCCCTGCAAGATTCAG 244
DB 1916 CTGGGCGGCGCCAGCCACGCTCAGCTGAGTATCCCGCGGTCGCGCCGAGCAGCG 1975
QY 245 GCACCTTCCTCGAGAGCCAGGAGTATATACAGGGGAGAGAAATCTCCAGAGAAAG 304
DB 1976 GCACATGATGTGGGAGTGAAGCCGCGGAGCCATGACAAAGCAGCTGCGCAGAGAGT 2035
QY 305 AAATTACAATCAGATCAGGAGCAGCACTACTCTCTCGAAACCTCAGTGATCACAGAG 364
DB 2036 ACCGTGCGGTCAGAGCCCTGAGAGCCCTCGCTCAGCAGAGAACTTGACCGACTCTCTGG 2095
QY 365 TGGCCATCAGAGTTCACCACTTTAGACTGTCTATGCTTAATGTTGTCCTCCGAGGCTCAGA 424
DB 2096 TGAACGTGAGCACTCGCTGGAGATGAGTGTGTTGGTGGCGGAGCCGAGCCGAGCAGCA 2155

QY 425 TCACITGGTTTAAACAAACACACAAATACACAGAGCCCTGGAAATATTTTAGGACCA 184
DB 2156 TCGTGTGGTACAAACACAGAGGCTGCTGGAGCAAAAGTCTGGAGTGCAGTTCGGGACT 2215
QY 485 GAAGCAGCAGCGTGTATTGAAGAGTACACAGAGAGATCAAGGTGTCTATCACTCA 544
DB 2216 CCAACAGAGCTGAGCAACAGCGGTGCGGAGAGAGATCGGGAGCACTCTGTGA 2275
QY 545 AAGCCACCAACACAGAGGCTGTGTGAAGTTCAGCATCTCACTGTTCAGAGAGCT 604
DB 2275 GCGTGTGCAACGCGCAAGGGCTGGGTCAAGTCTCGCCAGGCTGGCGTGAAGGCTCG 2335
QY 605 CGGACAGTCTAATCTGAGGTGATCACTTAACATGACACCTGTGTGGCTGGCACTCT 564
DB 2336 AGGTAAGGCGAGCAATGAGATCGTGNCTTTCGGTACCGGCTCATCGCTGTCTCT 2395
QY 665 TCTGGCTCTTATTAACCTCTTATCGGAAATGAAAAGTCTTC---TCTGAAATAA 721
DB 2396 TCTGGCT 2455
QY 722 AGACTGACTACCTATCAATTAATGAGCCAGATGAGTTCCTTTGGATGACAGTGTG 781
DB 2456 AGAGGGCTTACCTGTCCATCATGAGACCCGGGAGGTCCCTCTGGAGGACAAIGCG 2515
QY 782 AGCGGCTCCCTTATGATGCCAGAGTGGAGTTTCCCGGGAGAGACITTAACCTGGCA 841
DB 2516 AATACCTGTCTTACGATGCCAGCAGTGGGAATTCGCCAGAGCGGCTGCACCTTGGGA 2575
QY 842 AATCAGTGGAGAGGGCTTTTGGAAAGTGTTCAGCAATGAGCAITTTGGCAITAGA 901
DB 2576 GAGTGTGGGTACGGGCTTGGGAAGGTGGTGAAGGCTCGCTTTTGGCATCCACA 2635
QY 902 AATCAGTACGTCGGGACTGTGGCTGTGAAAATGCTCAAGAGGGGGCCAGCGCCAG 961
DB 2636 AGGCGAGCAGCTGTGACACCGTGCGCGTGAATGCTGAAGAGGGCGGCGAGCGCG 965
QY 962 AGTCAAGAGCTGTGAGTACGCTAATAATCTTGACCCAGCAITGGCCACATCTGAAC 1021
DB 2696 AGCAGCGCGCTGTGATGCGGAGCTCAAGATCTCAITTCAGATCGGCAACCACTCAAC 2755
QY 1022 TGGTAACTGTGGGAGCTGCACCAAGAGAGGGGCTCTGATGTTGATTTGAAT 1081
DB 2756 TGGTCACTCTCTGGGGGTGACCAAGCGCGAGGGGCGGCTTGTGGTGTGGAGT 2815
QY 1082 ACTGCAATATGGAATCTCTCACTACCTCAAGAGCAAAAGTGTATTTTCTCA 1141
DB 2816 TGTGAAGTACGCAACCTCTCACTCTGCGCGCAAGCGGAGCGCTTCAGCCCT 2875
QY 1142 ACAAGGATGCAGCACTACCAATGAGGCTTGAAGAGAAATATGAGCGAGGCTGGAG 1201
DB 2876 GCGCGAGAGTCTCCGA-----GCGCGGAGCGCTTCCCGGCAATGTTGGAGC 2926
QY 1202 AAGGCAAGAACCAAGACTAGATAGGCTACCAAGCAGCAAGGCTTGGAGCTCCGCT 1261
DB 2927 TCGGAGGCTGGATCGAGCGCGGAGGAGCAGGACAGGCTCTCTTCGCGGCTCT 2986
QY 1262 TTCAGAGACATAAAGTCTGAGTGTGTGAGGAGAGGAGGATCTGACGTTTCTACA 1321
DB 2987 CGAAGACGAGCGGCGAGCGGCGGCTTCTCCAGACCAAGAGCTGAGGACCTGTGGC 3046
QY 1322 AGGAGCCCATCACTATGGAAGATCTGATTTCTTACAGTTTCAAGTTCAGGCGAGGCTAG 1381
DB 3047 TGAGCCGCTGACCATGAGATCTGTCTGCTACAGCTTCCAGGTGGCGGAGGATCG 3106
QY 1382 AGTTCTGTCTCCAGAAAGTCAATTCAGGAGCTTGGAGGAGGAGGAGGAGGAGGAGG 1441
DB 3107 AGTTCTGGCTTCCGAAAGTGCATCCAGAGAGCTGGCTGCTCGAGATCTGCTGT 1466
QY 1442 CTGAGAAACGCTGGTGAAGATTTGTATTTTGGCTTGGCCGGGATATTAAGAGAGC 1501
DB 3167 CGGAAAGCGAGTGGTGAAGATCTGTGACTTTGGCTTGGCCGAGATCTACAAAGAGC 3226
QY 1502 CGGATATGTGAGAAAGAGGATGACTGAGTCTCTGTGAAATGAGATGCTCCGAGTCA 1561

DB 3227 CTGACTACGTCGCAAGGCGAGTGGCGGCTGCCCTGAGTGTGAGTGGCCCTGAAAGCA 3286
QY 1562 TCTTTCAGAAATCTACAGCAAGAGAGAGAGAGTGTGTCTTTACGGAGTATGCTGTGGG 1521
DB 3287 TCTTCAGAGAGTGTACACCAAGAGAGTGTGTGTCTTTGGGGTCTTCTCTGGG 3346
QY 1622 AAATCTCTCTCTTAGTGGTCTCCATACCAAGAGTACAAATGATGAGGAGCTTTGCA 1681
DB 3347 AGATCTCTCTCTGGGGGCTCCCGGTACCTTGGGTGAGATCAATGAGGAGTCTTGGC 3406
QY 1582 GTCCCTGAGGAGGCAATGAGAGAGAGTCTCTGAGTACTTCTTCTGAAATCTATC 1741
DB 3407 AGCGCTGAGAGAGGCAAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3466
QY 1742 AGATCACTGAGTCTGCGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1801
DB 3467 GCACTGAGTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3526
QY 1802 TGAAGAACTAGTGTGATTTGCTTCAAGCAAAATGTACACAGAGGATGGTAAAGA 1853
DB 3527 TGGAGATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3578

RESULT 5

US-09-282-610-45

: Sequence 45, Application US/09982610

: Patent No. US20020146420A1

: GENERAL INFORMATION:

: APPLICANT: Genentech, Inc.

: Bennett, Brian D.

: Goeddel, David

: Lee, James M.

: Matthews, William

: Tsai, Siao Ping

: Wood, William I.

: TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

: NUMBER OF SEQUENCES: 45

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Genentech, Inc.

: STREET: 460 Point San Bruno Blvd

: CITY: South San Francisco

: STATE: California

: COUNTRY: USA

: ZIP: 94080

: COMPUTER READABLE FORM:

: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: WinPatIn (Genentech)

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/982,610

: FILING DATE: 17-Oct-2001

: CLASSIFICATION: <Unknown>

: PRIORITY APPLICATION DATA:

: APPLICATION NUMBER: 08/446,648

: FILING DATE: 1995-MAY-23

: APPLICATION NUMBER: 08/222616

: FILING DATE: 04-APR-1994

: ATTORNEY/AGENT INFORMATION:

: NAME: Lee, Wendy M.

: REGISTRATION NUMBER: 40,378

: REFERENCE/DOCKET NUMBER: P0821P3PCT

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 415/225-1994

: TELEFAX: 415/952-9881

: TELEX: 910/371-7168

: INFORMATION FOR SEQ ID NO: 45:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 9108 base pairs

: TYPE: Nucleic Acid

: STRANDEDNESS: Single

: TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-982-610-45

Query Match 25.81; Score 541.6; DB 10; Length 3108;
Best Local Similarity 58.73; Pred. No. 7.2e-150;
Matches 981; Conservative 0; Mismatches 679; Indels 12; Gaps 2;

QY 185 CTAAAGAGACATCCATACATCTTAATCTTACCAATCAATGATGTTTCCCTGCNAGATTCAG 244
DB 2848 CTGGGCGCGCCACGCCACGCTACGCTGATATCCCGCGGTCCGCGCCGAGCAGGAG 2907
QY 245 GCACCTAAGCTTCAGAGCCAGGAAATGATACACAGGAGGAGAAATCTCCAGCAAGAAAG 304
DB 2908 GCCACTAATGCTGCGAAGTGCAGACCGCGCAGCCATGACAGACACTTGCACAGAGT 2967
QY 305 AATTAATACATACAGATCAGAGACACCATACCTCTCTGGAACTCAGTATACACAG 364
DB 2968 ACCTGTGGGTGAGCGCCCTGAGACCCCTCGGCTACGACAGAACTTGACCGACCTCTGG 3027
QY 365 TGGCATACAGAGTTCACACCACTTACACTGTGATCTAATGGTGTCCCGAGCCCTCAGA 424
DB 3028 TGAACGTGAGGACTCTGTGAGATGAGTCTGTGGCGGAGCGCACCGCCCGAGCA 3087
QY 425 TCACCTGGTTTAAACACACCAAAATACACAGAGCCCTGGAAATATTTTAAAGACAG 484
DB 3088 TCGTGTGTACAAAGACAGAGGCTGCTGGAGAAAGTCTGGAGTGCAGCTTGGGACT 3147
QY 485 GAACGACGAGCTGTTTATCAAGAGCTACACAGAGAGATGAAGGTGTCTATACATGCA 544
DB 3148 CCAACACAGAGCTGAGCAATCAGCGCGCTGGGAGGAGATGCGGAGCGGTATCTGTGCA 3207
QY 545 AAGCCACCAACAGAGGCTCTGTGAAAGTTACAGTACCTACCTGTTCAAGAGAACT 604
DB 3208 GCGTGTCAACGCCAAGGCTGCTCACTCACTCTCCGCCAGCGTGGCGGTGAGAGCTCGG 3267
QY 605 CGNCAAGTCTAATCTGAGCTGATCACTCTACATGCACTGTGTGTGTGTGTGTGTGTGT 664
DB 3268 AGATAAGGAGCAGCAATGAGATCTGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3327
QY 665 TCTGGTCTCTATTACCTCTCTATCGGAAATGAAGAGTCTTC---TCTCAATATA 721
DB 3328 TCTGGTCT 3387
QY 722 AGACTGACTACTATCAATTAATGAGCCAGCATGAAATTCCTTTGATGAGCAGTGTG 781
DB 3388 AGAGGGCTACCTGCTCAATCAATGAGCCCGGAGGTCCTCTGAGAGCAATGCG 3447
QY 782 ACCGGCTCTCTATCATCCAGCAGTGGAGTTTGGCCGGAGAGACTTAACTGGCA 841
DB 3448 AATACCTGTCTTACATGCCAGCAGTGGAAATTCCTCGAGAGCGCTGACCTGGGA 3507
QY 842 AATCACTTGGAGAGGGCTTTTGGAAAGTGTTCAGCATCAGCATTTGGCAATTAAGA 901
DB 3508 GAGTGTCTGGCTACGGCGCTTGGGAAGTGTGTGAGGCTCTGCTTTGCGATCCCA 3567
QY 902 AATCACTACGTGCGGAGCTGTGGCTGTGAATGCTGAAGAGGGGGCCAGCGCCAGCG 961
DB 3568 AGGGAGCAGCTGTGCACACCGCTGGCCCTGAAATGCTGAAAGAGGGGGCCAGCGCCAGCG 3627
QY 962 AGTACAAAGCTGTGAGTACTGAGTAAATCTTGACCAATTTGGCCACATCTGAGCG 1021
DB 3628 AGACCGCGGCTGTGAGTGTGGAGCTCAAGATCTTATCATGCGCAACCTTCAAG 3687
QY 1022 TGGTTAACTGTGGAGCGCTTACCAGCAAGAGGCGCTCTGATGGTGAATGTTGAAT 1081
DB 3688 TGGTCAACCTCTCTGGGCGGTGACCAAGCGCGAGCGCCCTCATGGTGAATGAGTGT 3747
QY 1082 ACTGCAATATGGAATCTTCAACTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
DB 3748 TCTGCAAGTACGCAAGCTTCTCAACTTCTTCTGCGGCCCAAGCGGAGCGCTTCAAGCCCT 3807
QY 1142 ACAAGGATGCAGCTACATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201

DB 3808 GCGCGAGAGATCTCCCGA-----GCAGCGCGAGCGCTTCGCGCCATGTTGGAGC 3858
QY 1202 AAGCCAGAAACCAAGACTAGATAGCTGCACAGCAGCGCAAGAGCTTTTCGAGCTCCGGCT 1261
DB 3859 TCGCAGGCTGATCGAGCGCGCGGAGCAGCAGAGGTCTCTTTCGCGGGTCT 3918
QY 1262 TTAGAGAGATTAAGATCTGATGATGTTGAGGAGAGAGAGATTTCTGACGGTTTCTACA 1321
DB 3919 CGAAGACCGAGGCGGAGCGAGCGGCGCTCTCCAGACCAAGAGCTGAGGACTGTGGC 3978
QY 1322 AGAGCGCCATACATATGAAGATCTGATTTTACAGTTTTCAGTGTGCGGAGGATG 1381
DB 3979 TGAGCGCGCTGACCAATGAGAGATCTTCTGTACAGCTTCCAGTGTGCGGAGGATG 4038
QY 1382 AGTTCCTGTCTTCCAGAAATGCTCATTCGAGCTTGCAGCGAGAGAAATCTTTTAT 1441
DB 4039 AGTTCCTGTCTTCCAGAAATGCTCATTCGAGCTTGCAGTGTGCGGAGGATG 4092
QY 1442 CTGAGACAGCTGTGAGATTTGTGATTTTGGCTTGCCTGGGATTTTAAAGAC 1501
DB 4099 CGAAGCGAGCTGTGAGATCTGTGATTTTGGCTTGCCTGGGATTTTAAAGAC 4158
QY 1502 CCGATATGTGAGAAAGAGATACTCGACTTCTCTGAATGAGTGCCTCCGAAATCTA 1561
DB 4159 CTGACTAGCTTCCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4218
QY 1562 TCTTGACAAATCTTACAGCAGCAGAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1621
DB 4219 TCTTGACAAATCTTACAGCAGCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4278
QY 1522 AATCTCTCTTGT 1681
DB 4279 AGATCTCTCTGTGGGCGCTTCCCGGTACCTGGGTGAGATCAATGAGGATGTGTGTGT 4339
QY 1582 GTGCGCTGAGGAGAGCAGATGAGATGAGACTTCTCTGACTTCTACTCTCAATCTATC 1711
DB 4339 AGCGCTGAGAGAGCAGCAGAGATGAGGCGCGCGGAGCTGCGGAGCTGCGGAGCTGCGG 4398
QY 1742 AGATCATCTGAGCTTCTGCGCAGAGACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1801
DB 4399 GATCATCTGAGCTTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4458
QY 1802 TGGAAAGCTAGTGTGATTTGTTCAAGCAATGTACACAGATGTTAAAGA 1853
DB 4459 TGGAGATCTGCGGAGAGCTGTCTCCAGGCGAGGCGCTTCAAGAGAGAGAGAGAG 4510

RESULT 6

US-09-982-610-17/c

: Sequence 17, Application US/09982610

: Patent No. US2002014620A1

: GENERAL INFORMATION:

: APPLICANT: Genentech, Inc.

: Bennett, Brian D.

: Goeddel, David

: Lee, James M.

: Matthews, William

: Tsai, Siao ping

: Wood, William J.

: TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

: NUMBER OF SEQUENCES: 45

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Genentech, Inc.

: STREET: 460 Point San Bruno Blvd

: CITY: South San Francisco

: STATE: California

: COUNTRY: USA

: ZIP: 94080

: COMPUTER READABLE FORM:

: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: WinPatIn (Genentech)

Db 2343 CCAGGTGGCAAGGCGATGGCTTTCTCCCTCCAGAAATGATTCACAGACAGACTGGC 2402
Qy 1422 AGCGAARACATCTTTTATCTGAGAACACAGTGTGAGATTTGTGATTTGGCCTTGC 1481
Db 2403 AGCGAARACATCTTTTATCTGAGAACACAGTGTGAGATTTGTGATTTGGCCTTGC 2462
Qy 1482 CCGGATATTTATAGAACCCCGGATTAATGATGAGAAAGGAGATATCTGATCTGAA 1541
Db 2463 CAGAGACATCAGATGATTTTATGATGATGATGATGATGATGATGATGATGATG 2522
Qy 1542 ATGGATGGTCCGATCTATCTTTCAGAAATCTACAGACACAGAGAGAGAGAGAG 1601
Db 2523 GTGGATGGACCTTGAAGCAATTTTCACTTGTATACAGCTTGAAGTGGTGC 2582
Qy 1602 TTACGGATTTCTGTGGGAATCTTCTTGTGAGAGAGAGAGAGAGAGAGAGAG 1661
Db 2583 CTATGGATTTCTGTGGAGCTTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2642
Qy 1662 AATGGATGAGGACTTTTTCAGTCCCTGAGGAGAGAGAGAGAGAGAGAGAGAG 1721
Db 2643 GTTCGATTTCTAGTTCTACAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAG 2702
Qy 1722 CTCTATCTGAGTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1781
Db 2703 CGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2762
Qy 1782 GCCAAGATTTGAGAACTTTGAGAAACTAGTGTGATTTGCTTCAAG 1828
Db 2763 ACCAATTCAGCAATTTGTCAGTATTTGAGAGAGAGAGAGAGAGAGAGAG 2889

RESULT 8
US-09-967-768A-277
; Sequence 277, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: Patent version 3.0
; SEQ ID NO: 277
; LENGTH: 5084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-277

Query Match 12.3%; Score 259; DB 10; Length 5084;
Best Local Similarity 55.8%; Pred. No. 3.4e-55;
Matches 595; Conservative 0; Mismatches 440; Indels 32; Gaps 4;
Qy 790 CTTTATGATGACACAGTGGAGTTGGCCGGGAGAGACATTAACCTGGCAATACAT 849
Db 1747 CTCTTATGATCAAAATGGAGTTTCCAGAAACAGCTGAGTTTGGGAAACCCCTG 1806
Qy 850 GGAAGAGGGCTTTTGGAAAGTGGTTCAAGCATCAGCATTTGTCATTAAGAAATACCT 909
Db 1807 GTGTGGAGCTTTCCGAGAGTTTGTGAGGCACTGTTAAGCTTATTAAGTACAT 1866
Qy 910 ACGTGGGAGCTTGGCTGTGAAATGCTCAAGAGAGGGCCAGCGCCAGAGTACAAA 969
Db 1867 CGCGGATGATGCTGCTGTAAAGATGCTCAAGCGAGGCGCCATTTGAGAGAGAGGNA 1926
Qy 970 GCTCTGATGATGAGCTAAATAATCTTGACCCGACATTTGCCACCATCTGAACGTGTAAAC 1029

Db 1927 GCCTCATGTCTGAACCTCAAAAGTCTCTGAGTTACCTTGGTATCATCATGATTTGTGAAT 1986
Qy 1030 CTCTGGGAGCTTGCACCAAGAGAGAGGCTCTCTGATGGTGAATTTGTGATGATGATGAT 1099
Db 1987 CTACTTGGAGCTTGCACCA---TTGGAGGCGCCACCTGGTCAITACAGAAATATTGTTC 2048
Qy 1090 TATGGAATCTCTCAACTACCTCAAGAGCAACAGCTGACTTATTTTCTCAACAAG--- 1156
Db 2044 TATGGTGAATCTTTGAATTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2104
Qy 1147 -----GATCGAGCACTACACATGAGCCCTAGAGAGAGAGAGAGAGAGAGAGAG 1204
Db 2104 GAGATCATCTGAGAGAGCTGACATTTATAGAAATCTCTGATTCATTCAGAGAGAGTCTTCCCTG 2163
Qy 1195 CTGGAACAAGGCAAGAAACCAAGACTAGATAGCTGCTACAGCAGCAGAGAGAGAGAGAG 1263
Db 2163 CAGCGATAGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2222
Qy 1255 TCGGCTTTTCAAGAGATTAAGTCTGATGATGATGATGATGATGATGATGATGATGATG 1314
Db 2223 CAAGGCGGCAAG 2282
Qy 1315 TTT-----TACAAGGAGCGCCATCATCTATGAGAGATCTGATTTCTTACAGATTT 1374
Db 2283 CGCATCATCTGAG 2442
Qy 1362 TCAAGTGGCCAG 1421
Db 2343 CCAGGTGGCAAGGCGATGGCTTTCTCGGCTCCCAAGATTTGATTCACAGAGAGAGAG 2402
Qy 1422 AGCGAARACATCTTTTATCTGAGAACACAGTGTGAGATTTGTGATTTGGCCTTGC 1481
Db 2403 AGCGAARACATCTTTTATCTGAGAACACAGTGTGAGATTTGTGATTTGGCCTTGC 2462
Qy 1482 CCGGATATTTATAGAACCCCGGATTAATGATGAGAAAGGAGATATCTGATCTCTGAA 1541
Db 2463 CAGAGACATCAGATGATTTTATGATGATGATGATGATGATGATGATGATGATGATG 2522
Qy 1542 ATGGATGGTCCGATCTATCTTTCAGAAATCTACAGACACAGAGAGAGAGAGAGAGAG 1601
Db 2523 GTGGATGGACCTTGAAGCAATTTTCACTTGTATACAGCTTGAAGTGGTGC 2582
Qy 1602 TTACGGATTTCTGTGGGAATCTTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1661
Db 2583 CTATGGATTTCTGTGGAGCTTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2642
Qy 1662 AATGGATGAGGACTTTTTCAGTCCCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1721
Db 2643 GTTCGATTTCTAGTTCTACAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2702
Qy 1722 CTCTATCTGAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1781
Db 2703 CGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2762
Qy 1782 GCCAAGATTTGAGAACTTTGAGAAACTAGTGTGATTTGCTTCAAG 1828
Db 2763 ACCAATTCAGCAATTTGTCAGTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2889

RESULT 9
US-09-967-768A-277
; Sequence 3, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ibor R.
; TITLE OF INVENTION: TROPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: IW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3035
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-919-408-3

Query Match 11.9%; Score 249.4; DB 10; Length 3501;
Best Local Similarity 55.9%; Pred. No. 1.9e-63;
Matches 572; Conservative 0; Mismatches 421; Indels 30; Gaps 4;

QY 792 TATGATGCCAGCAAGTGGAGTTGGCCGGAGAGACATAACTGGGCAAAACACTTGG 851
Db 1848 TGAATATGATCTCAAAATGGAGTTTCCAGAGAGAAATTTAGAGTTTGGCAAGTACTAGG 1907
QY 852 AAGAGGGCTTTTGGAAAGTGGTTCAGCAATCAGCAATTTGGCATTAAGAAATCACCTAC 911
Db 1908 ATCAGGTGCTTTTGGAAAGTGGTTCAGCAATCAGCAATTTGGCATTAAGAAATCACCTAC 911
QY 912 GTGCCGCTGTGGCTGTGAAAATGCTTGAAGAGGGGGCCACGGCCAGCGATACAAAGC 971

Db 1958 CTCAATCCAGGTTGCCGTCAAAAATGCTGAAAGAAAGACAGACAGCTCTGAAAGAGAGGC 2027
QY 972 TCTGATGACTGAGCTAAATAATCTTGACCCACATTTGCCACATCTGACGTTGTTTACCT 1031
Db 2028 ACTCATGTGAGACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTTGTAACCT 2097
QY 1032 CTCTGGAGCCTGCACCAGCAAGAGAGGCTCTCTGATGTTGTTGATTTGATTTGCAATA 1091
Db 2086 CTCTGGGGGCTGCACACTGTGAGG--CCAATTTACTTTGATTTTGAATGTTGCTA 2144
QY 1092 TGGAAATCTCTCAACTACCTCAAGCAAAAGCTG-----ACT 1129
Db 2145 TGGTATCTCTCAACTATTAAGAAAGTAAAGAGAAATTTTACAGGACTTGGACAGA 2204
QY 1130 TATTTTCTCAACAAGAGTGCAGCACTACATGAGCTTAAGAAAGAAATAATGAGC 1189
Db 2205 GATTTTCAAGGAPACAAATTTTACCTTTTACCCCACTTTTCCAATCACAATTCAG 2264
QY 1190 CAGCCTTGGAAACAGCAAGAACCAAGCTAGATAGCTTACCAAGAGGAAAGCTT-- 1247
Db 2265 CATGCTTGGTTCAGAGAAAGTTTCAATACACCGGACTCGGATCAAAATCTCAGGGCTTCA 2324
QY 1248 TCCGAGCTCCGGCTTTTCAAGAAAGTCTGAGTGATGTTGAGGAGAGGAGGATTC 1307
Db 2325 TCGGAATTCATTTTCACTCTGAGATGAATTTGAATATGAACCAAAAGAGCTTGAAGA 2384
QY 1308 TGACCTTTTCTCAAGAGGACCACTATGAGAGATCTGATTTTCTTACAGTTTCAAGT 1367
Db 2385 AGAGGAGGCTTCAATGTGC---TTCAATTTGAAGATCTTCTTCTTGCATATCAAGT 2441
QY 1368 GGCAGAGCAAGTGGTCTGCTTCCAGAAAGTCAATTCATCGGACCTGGCAGGGAG 1427
Db 2442 TCCAAAGGAATGGAATTTCTGGAAATTTAAGTCTGTGTTTACAGAGACCTTGCOCGCA 2501
QY 1428 AAACATTTCTTTTATCTGAGACACAGCTGTGAAGATTTGTGATTTTGGCTTGCCTGGGA 1487
Db 2502 GAAGCTGCTTGTCCACCCAGGGAAGTGGTGAAGATATGTGATTTGGATTGGCTCGAGA 2561
QY 1488 TATTTAAGACCCCGATTTGTGAGAAAGGAGATCTCGACTTCTCTGAAATGGAT 1547
Db 2562 TATCATGATGATTTCCAACTATTTGTGAGGCAATGCGCTCTGCTTAAATGGAT 2621
QY 1548 GGCTCCGGAATCTATCTTTGACAAATCTACAGCACCAAGAGGAGCTGTGCTTACGG 1607
Db 2622 GGCCCGGAAAGCTTTTGAAGGCACTTACACCATTAAGATGATGTTGGTCAATGG 2681
QY 1608 AGTATCTGTGGAAATCTTCTCTTGTAGTGGTCTCCATACCCAGAGTACAAATGGA 1667
Db 2682 AATAATTACTGTGGAAATCTTCTCTTGTGTAATCTTACCTTGGCATTCGGGTGGA 2741
QY 1558 TGAGGACTTTTGCAGTCCCTGAGGAGGCAATGAGGATGAGGCTCTCTGATCTTAC 1727
Db 2742 TGTAACTTCTCAAACTGATTCAAATTTGAATTTAATGATGATGATGATGATGATGAT 2801
QY 1728 TCTTGAATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1797
Db 2802 AGAAGAAATATACATTAATGCAATCTCTGCGGCTTTTGTACTCAAGGAAACGGCATC 2861
QY 1788 ATT 1790
Db 2862 CTT 2864

RESULT 10
US-09-872-136-3
: Sequence 3, Application US/09872136
: Patent No. US20020119545A1
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ibor R.
: TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
: RECEPTORS AND THEIR LIGANDS
: NUMBER OF SEQUENCES: 10

TYPE: DNA
ORGANISM: Homo sapiens
US-09-944-807-9

Query Match 11.1%; Score 233.6; DB 10; Length 3992;
Best Local Similarity: 54.3%; Pred. No. 1e-58;
Matches 539; Conservative 0; Mismatches 444; Indels 9; Gaps 3;

QY 805 AAGTGGGAGTTCCTGGGAGAGCTTAACCTGGCAATACATCTGGAGAGGGCTTTT 864
|||||
DB 2020 AAGTGGGAGTTCCTGGGAGAGCTTAACCTGGCAATACATCTGGAGAGGGCTTTT 2079
|||||

QY 865 GGAAGAGTGGTTCAGCATACATCTGGCAATTAAGAAATACCTACCTGCGGACTGTG 924
|||||
DB 2080 GGAAGTGGTTCAGCATACATCTGGCAATTAAGAAATACCTACCTGCGGACTGTG 2139
|||||

QY 925 CTGTGAAATCTGAAAGAGGGGCGCAGCGCAGGAGTCAAAAGCTCTGATGACTGAG 984
|||||
DB 2140 CTGTGAAATCTGAAAGAGGGGCGCAGCGCAGGAGTCAAAAGCTCTGATGACTGAG 2199
|||||

QY 985 CTAAATCTGTGACCCACATCTGACCATCTGACCTGAGCTGTTAACTCTGGAGCTGC 1044
|||||
DB 2200 CTGAGATCAATGAGCTGACCTGCGCAGCAGCAGCAATCTGCACTCTTGGAGCTGT 2259
|||||

QY 1045 ACCAAGCAGAGGCGCTCTGATGATGTTGAAATCTGCAAAATATCGAAATCTCTCC 1104
|||||
DB 2260 ACC---CATGAGGCGCTCTGATGATGTTGAAATCTGCAAAATATCGAAATCTCTCC 2316
|||||

QY 1105 AACTACTCTAAGAGCAAGCTGA---CTTATTTTCTCAAGAGATGCGAGCTACAC 1161
|||||
DB 2317 AACTTTCTGGAAGAGGCTGAGGCTATGCTGGGACCGAGCTGAGCGCGGCGGAGAC 2375
|||||

QY 1162 ATGAGCTTAAGAAAGAAATAGAGCCAGGCTTGAACAAGGCAAGAACCAAGACTA 1221
|||||
DB 2377 CCCGAGGAGGCTGACTATAAGCAATCACTGCGAGAGAAATATGTCGCGAGGAGC 2436
|||||

QY 1222 CATAGCTCATCAGCAGGAGAGCTTTGGAGCTCCGGCT---TCAGGAAGATAAAGT 1278
|||||
DB 2437 AGTGGCTCTCCAGCTAGGCTGAGCAGCTATGTGAGATGAGGCTGTCTCCTCTCT 2495
|||||

QY 1279 CTGAGTATGTTGAGAGAGAGGATCTGACGGTTTCTCAAGAGGAGGCTCACTATG 1338
|||||
DB 2497 TCAATGACTCTCTCTGACAGAGCTGCAAGAGGATGAGCGGCTCTGGAGCTC 2556
|||||

QY 1339 GAGATCTGATTTCTACAGTTTCAAGTGGCCAGAGGAGTGGAGTCTCTGTTCTCAGA 1398
|||||
DB 2557 CGGAGCTGCTCTCTCTGACAGAGTATAGCCAGGAGTGGCTCTCTGCTCTCAG 2616
|||||

QY 1399 AAGTGAATTCATCGGAGCTGCGAGCAGAGAACTCTTTATCTGAGACAGAGCTGGTG 1458
|||||
DB 2617 ANTGTATCCAGGAGGCTGCGAGCAGGCTGAGGCTGTTGACCAATGTTGAGGCT 2676
|||||

QY 1459 AAGATTGTGATTTGGCTTCCCGGGATATTTAAGACCCCGATTATGTGAGAAAT 1518
|||||
DB 2677 AAGATTGGGAGCTCGGCTGCTAGGAGCATCATGATGACTCACTACATGTTCAAG 2736
|||||

QY 1519 GGAGATCTGACTCTCTCTGAAATGAGTGGCTCGGAACTATCTTTTGACAAATCTAC 1578
|||||
DB 2737 GGCAATGCGGCTCTCTCTGAAATGAGTGGCTCGGAACTATCTTTTGACAAATCTAC 2796
|||||

QY 1579 AGCAACCAAGAGGAGTGTGCTTACGAGTATTCCTGTGGAACTCTCTCTTAGGT 1638
|||||
DB 2797 ACGGTTACAGAGAGCTGTGCTTACGAGTATTCCTGTGGAACTCTCTCTTAGGT 2856
|||||

QY 1639 GGGTCTCATACCCAGGAGTACAAATGATGAGGACTTTTCAGTGGCTGAGGAGAGGC 1698
|||||
DB 2857 CTGAATCCCTACCTGCGATCTCTGAGCAGCAGAGTCTATTAAGTGGTGAAGGATGA 2916
|||||

QY 1699 ATGAGATGAGCTCTCTGAGTCTCTCTCTGAAATCTATCAATCATCTGCTGACTGC 1758
|||||
DB 2917 TACCAATGCGGCTCTCTCTGAGTCTCTCTCTGAAATCTATCAATCATCTGCTGACTGC 2975
|||||

QY 1759 TGGCAGAGAGAGCCCAAAAGAGAGCCCAAGAT 1790
|||||

Db 2977 TGGGCTTGGAGCCACCCACAGACACCCACCTT 3008
|||||

RESUL 13
US-09-919-408-1
Sequence 1: Application US/09919408
Patent No. US20020072077A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 112..3006
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 31..111
FEATURE:
NAME/KEY: CDS

LOCATION: 31..3009
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-919-408-1

Query Match 10.5%; Score 221.6; DB 10; Length 3453;
Best Local Similarity 53.0%; Pred. No. 3 3e-55;
Matches 585; Conservative 0; Mismatches 489; Indels 30; Gaps 4;

QY 785 GGCCTCCCTATATGATCCAGCAAGTGGAGTTTGGCCGGGAGAGACITTAACATGGSCAAAT 844
DB 1817 GGGACTATGAATAIGACCTTAATGGGATTCGCCGAGAGAACTTAGAGTTTGGGAAG 1876

QY 845 CACITGGAAGAGGCGCTTTTGAAGTGGTTCAGCAACAGCAITTCGCAITTAAGAAAT 904
DB 1877 TCTGGGCTTCGCGCTTTTCGGAGGGGTATGAACGCCAGCGCTTAGCATTAGTAAAT 1936

QY 905 CACCTACGTCGCGAGCTTGGCTGTGAAATGCTCAAGAGGGGCGCCAGCGCCACCGAGT 964
DB 1937 CGGGAGTCTCAATTCAGGTGGCGGTAGAGTGTAAAGAGAAAGCTGACAGCTGAGAA 1996

QY 965 ACAAGCTCTGATGACTGACATAAATTCATGACCCACATTCGCCACATCTGAACGTGG 1024
DB 1997 AGAAGCTCTGATGCGGATCAAAATGATGACCCACCTGGGACACCAATGACAAATCG 2056

QY 1025 TTAACCTCTGGGAGCGCTGCAACCAAGAGGAGGCGCTCTGATGGTGTGTTGTAATCT 1084
DB 2057 TGAATCTGCTGGGCAAGCAC---ACTGTAGGGCCAGTGTACTTGATTTTGAATAT 2113

QY 1085 GCAATATGGAATCTTCCAGCTACCTCAAGAGCAAAACGCT----- 1126
DB 2114 GTTGTATGGTGGCTCTCACTACCTAAGAGTAAAGAGAGAAAGTTTACAGGACAT 2173

QY 1127 ---ACTTATTTTCTACAGAGTACGAGTACATGAGGCTTAAGAGAGAAAT 1182
DB 2174 GACAGAGATTTTAAAGAAATATTTTCACTGTCTTACCTCTTCCAGGCAATCAA 2233

QY 1183 ATGAGCCAGCGCTTGAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1242
DB 2234 ATCCAGCATGCTGTGTACAGAGAGTTCAGTTACACCCGCTTGGATCAGCTCAG 2293

QY 1243 AG---CTTTCAGGCTCGGCTTTCAGAGAGATAAAGTCTGAGTGTGAGGAAGA-- 1298
DB 2294 GGTCAATGGGAATTCATTAATCTGAGAGATGAGATTAATTAAGAAACCAAGAGAGG 2353

QY 1299 -GGAGGATCTGAGCGTTTCTACAGAGGAGCCATCACTATGGAAGATCTGATTCITACA 1357
DB 2354 TGGCAGAGAGAGAGAGAGATTTGACGTCTGACGTTTGGAGACCTTCTTCTTCTG 2413

QY 1358 GTTTCAGTGGCCAGAGGATGAGTCTCTGCTTCCAGAAAGTGCATTCAGCGGAC 1417
DB 2414 CGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTGTCCACAGAGAC 2473

QY 1418 TGGCAGCAGAAACATTTTATCTGAGAACACGTTGGTGAAGATTTGATTTTGGCC 1477
DB 2474 TGGCAGCAGAGATGTTGTTGTCACCCAGGAGGTTGTAAGATCTGTGACCTTGGAC 2533

QY 1478 TTGCCCGGATATTTAAGAACCCCGATTATGTGAGAAAGAGAGTCTGACTTCTC 1537
DB 2534 TGGCCCGAGACATCTGAGCGACTCCAGTCTGTCAGGGGCAACGCGCTGGCGG 2593

QY 1538 TGAATGATGCTCGGATCTATCTTTCAGAAATCTACAGCAGCAGAGAGAGCTGT 1597
DB 2594 TGAATGATGCTCGGAGCTATTTTGAAGGATCTACAAATCAAGATGAGCTGT 2653

QY 1598 GGTCTTACGAGTATCTGTGGAAATCTTCTCTTGTAGGTGGTCTCCATACCCAGAG 1657
DB 2654 GGTCTTACGAGTATCTGTGGAGATTTTCTGCTGGGTGAGACCTTACCTTGGCA 2713

QY 1658 TCAATGATGAGGATTTTGCAGTCCCTTCAAGGAGGAGCATGAGGATGAGCTCTG 1717
DB 2714 TTCTGTGCGAGCTTATCTTATAAATGATTCAGAGTGGATTTAAATGAGAGGACAT 2773

QY 1718 AGTCTCTACTCTGAAATCTATCATGATCTCTGAGTCTCTGAGCAAGAGCAAG 1777

DB 2774 TCTATGCCACAGAGGATATACITTTGTAATGCAATCTCTGGCTTTGACTCAGGA 2833
QY 1778 AAAGCCCAAGATTTGAGAACTTGGAAACACTAGTGATTTCTTCAAGCAAAATGTAC 1837
DB 2834 ACGGCCCACTCTTCCCACTGACTTATTTAGGATGTCAGCTGCGAGAGGAGCAAG 2893
QY 1838 AACAGGATGTTAAAGACTATACCTC 1861
DB 2894 AAGCATGATCAGAACTCATCTC 2917

RESULT 14
US-09-872-136-1
Sequence 1, Application US/09872136
Patent No. US20020119545A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872.136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208.786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021.324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977.451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906.397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813.593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793.065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728.913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679.666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28.601
REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 112..3006

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 31..111

FEATURE:

NAME/KEY: CDS

LOCATION: 31..3009

SEQUENCE DESCRIPTION: SEQ ID NO: 1;

US-09-872-136-1

Query Match 10.5%; Score 221.6; DB 10; Length 3453;

Best Local Similarity 53.0%; Pred. No. 3.3e-55;

Matches 585; Conservative 0; Mismatches 489; Indels 30; Gaps 4;

785 GGCCTCCCTTATGATGCCAGCAAGTGGAGTTTGGCGGGAGAGCTTAACTGGGCAAT 844

1817 GGGACTATGATATGACCTTATGCTGGAGTTCCGAGAGAGACTTAGAGTTGGGAGG 1876

845 CACTTGGAGAGGGGCTTTTGGAAAAGTGGTTCAAGCATCAGCAITTTGGCATTAAGAAAT 904

1877 TCCTGGGCTCTGGCGCTTTGGGAGGGTGTGACGCCAGCGCTATGGCAITAGTAAAT 1936

905 CACCTAGTCCCGAGCTGTGGCTGTGAATGTGAAGAGGGGCCAGGCCAGCGAGT 964

1937 CGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAAGCTCAGACGTGTGAAA 1996

965 ACAAGCTCTGATGACTGAGCTTAAATATTTGACCCACATTTGCCACCATCTGAACGTG 1024

1997 AAGAGCTCTCATGTGGAGCTCAATGATGACCTTGGCAGCAGCAATGCAACATCG 2056

1025 TPACTTGTGGAGCTGTGACCAAGAGAGGGGCTCTGATGTGATTTGATACAT 1084

2057 TGAATCTGTGGGGGATGACAC---ACTGTGAGGGGAGTGTGATTTTGAATAT 2113

1085 CCAATATGAATCTCTCACTACCTCAAGAGCAAGCTG----- 1126

2114 GTTCTATGTTGACCTCTCACTACCTCAAGAGCAAGCTG----- 1126

1127 ----ACTATTTTTCACAAAGAGTCCAGCAGCTACATGGAGCCCTTAAGAAAGAAA 1182

2174 GGACAGAGATTTTAAAGAACATAATTCAGTTCTTACCTTACCTTCCAGCAGCATCAA 2233

1183 ATGGAGCCAGGCTGTGAACAGGCAAGAACCAAGACTAGATAGGTACACAGCAGGAA 1242

2234 ATTCAGCATGCTGTGTTTACAGAGAAAGTTCAGTTACACCGCCCTTGGATCAGCTCAG 2293

1243 AG---CTTTGGAGCTCCGCGCTTTCAGGAAGTAAAGTGTGATGATGTTGAGGAAGA-- 1298

2294 GGTTCATGGAATCAATTCATCTGAGATGAGATGATGATGATGATGATGATGATGATG 2353

1299 -GGAGATCTGACGGTTTTCACAGGAGCCCATCATTGGAAGATCTGATTTTCTTACA 1357

2354 TGGCAGAAAGAGAGGAGGAAGTTTGAAGCTGTGAGCTTTGAGACCTCTTCTGCTTTG 2413

1358 GTTTCAAGTGGCAGAGGAGATGATTTCTGCTTCCAGAAAGTGTGATTCATCGGAGCC 1417

2414 CGTACCAAGTGGCCAAAGGAGTAAATCTCTGGAGTTCAAGTCTGTGTGTCACAGAGACC 2473

1418 TGGCAGCAGAAACATCTTTTATCTGAGAACACGTTGTGAGATTTGTGATTTTGGCC 1477

2474 TGGCAGCAGGAAATGTGTTGGTACCCAGCGGAGGTGTGAGATCTGTGATTTTGGAC 2533

1478 TTGCCCGGGAATTTATAGAACCCCGATTTGTTGAGAAAGGCAATCTGATCTGCTTCCCT 1537

2534 TGGCCCGGAGACATCTGTAGGAGCTCCAGCTACCTGTGTCAGGGGCAACGACGCTCGCG 2593

1538 TGAATGGATGGCTCCCGAATCTATCTTTGACAAAATCTTACAGTACCAAGAGGCGATGT 1597

Db 2594 TGAAGTGGATGGCAGCCGAGAGCTTATTTGAAGGATCTACACAATCAAGAGTGCCTCT 2653

QY 1598 GGTCTTACGGAGTATGCTGTGGAAATCTTCTCTTCTAGTGGTCTCCATACCCAGAG 1657

Db 2654 GGTCTTACGGAGTATGCTGTGGAGATATTTTCACTGGGTGTGAACCTTACCTGGCA 2713

QY 1658 TACAATGGATGAGGACTTTTTCAGTCTGCTGAGGAGGAGGAGGATGAGAGTCTG 1717

Db 2714 TTTCTGTGAGGCTTAACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAGCAGCA 2773

QY 1718 AGTACTCTAGTCTGAAATCTATCATGATCTGTGGAGTCTGTGGCAGAGAGCCCAAG 1777

Db 2774 TCTATGCCAGAGAGGATATCTTTTAAATGCAATCTCTCTGGGCTTTTGAATCAAGGA 2833

QY 1778 AAAGCCCAAGATTTGCAAGACTTTTGGAAAACCTAGTGTGATTTGTTCAACAATATGAC 1837

Db 2834 AGCGCCCACTTCCCAACCTGACCTTATTTAGGATGTCAGCTGGCAGGAGCAGCAAG 2893

QY 1838 AACAGGATGTTAAAGACTACATCC 1861

Db 2894 AAGCATGATCAGAACATCCATCC 2917

RESULT 15

US-09-866-510-1

: Sequence 1, Application US/09866510

: Patent No. US20020111304A1

: GENERAL INFORMATION:

: APPLICANT: KAZLAUSKAS, ANDRIUS

: APPLICANT: IKUNO, YASUSHI

: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES

: FILE REFERENCE: ERM-104.01

: CURRENT APPLICATION NUMBER: US/09/866,510

: PRIOR FILING DATE: 2001-05-25

: PRIOR APPLICATION NUMBER: 60/250,747

: PRIOR FILING DATE: 2000-12-01

: PRIOR APPLICATION NUMBER: 60/289,103

: PRIOR FILING DATE: 2001-05-07

: NUMBER OF SEQ ID NOS: 33

: SOFTWARE: PatentIn Ver. 2.1

: SEQ ID NO 1

: LENGTH: 3270

: TYPE: DNA

: ORGANISM: Homo sapiens

: FEATURE:

: NAME/KEY: CDS

: LOCATION: (1)..(3257)

: US-09-866-510-1

Query Match

Best Local Similarity 10.2%; Score 213.4; DB 10; Length 3270;

Matches 521; Conservative 0; Mismatches 486; Indels 96; Gaps 2;

QY 744 AATGACCCAGATGAAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCCCTTATGATCCAG 803

Db 1592 AATCAGCCCGATGAGCATGATATATTTATGTTGGACCGATGACGTGCTTATGACTC 1751

QY 804 CAAGTGGAGTTTCCCGGAGAGACTTAACTTGGCAATCACTTGGAGAGGGCTTT 863

Db 1752 AAGATGGAGTTTCCAAAGATGAGATGCTTGTGGGTCTTGGGGTCTGAGCGTT 1811

QY 864 TGGAAAGTGGTTTCAGCATCAGCATTTGCAATTTAGAAATCACCCTAGTGTGCGGACTGT 923

Db 1812 TGGAGGTGGTTTGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1871

QY 924 GGTGTGAAAATGCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 983

Db 1872 TGCAGTGAAGTGTAAATCAACCCAGGCGCAGATCCAGTGAACAGCTCTCACTGCTGA 1931

QY 984 GCTAAATATCTTGAACCCACATTTGGCCACCATCTGACGTGTGTTAACTTGTGGAGCCTG 1043

Db 1932 ACTGAAGATAATGACTTACCTTGGGGCCACATTTGAACATTTGTAACATTTGTTGGAGCCTG 1991


```

Db 601 CAGGAGCTGTGATGTCACGATATGTAATGCTTCAAGTTCATGAGCCIGGAAGA 666
QY 2023 ATCAAAACCTTTGAGAACTTTTACCGAATCCACCTCCATGTTGATGATACAGGCC 2062
Db 661 ATCAAAACCTTTGAGAACTTTTACCGAATCCACCTCCATGTTGATGATACAGGCC 720
QY 2083 GACAGCAGCACTCTGTGG 2101
Db 721 GACAGCAGCACTCTGTGG 739

RESULT 3
BQ604389/c
LOCUS BQ604389 697 bp mRNA linear EST 24-JUN-2002
DEFINITION MI-P-CP1-nzb-m-17-0-UI-s1 MI-P-CP1 Sus scrofa cDNA clone
AUTHORS MI-P-CP1-nzb-m-17-0-UI 3', mRNA sequence.
VERSION BQ604389
KEYWORDS EST.
SOURCE BQ604389.1 GI:21551115
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 697)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Tuggle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktuggle@iastate.edu
Tissue Procurement: Dr. Chris Tuggle, Iowa State University
cDNA Library preparation by: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA-yes.
Location/Qualifiers
1..697
/organism="Sus scrofa"
/strain="crossbreed"
/db_xref="taxon:9823"
/clone_lib="MI-P-CP1-nzb-m-17-0-UI"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not 1; Site 2: EcoRI. The MI-P-CP1
library is normalized library derived from the MI-P-CP0
library, ultimately derived from uterus tissue. For a
detailed description of the library from which this clone
was derived, please visit our web site at
http://pigest.genome.iastate.edu/. The procedure used to
create this library has been previously described (Bonaldo
Lennon and Soares, Genome Research 6: 791-806, 1996)
TAG_LIB=MI-P-CP1
TAG_TISSUE=uterus
TAG_SEQ=AGTCCAATCG"
BASE COUNT 139 a 181 c 181 g 136 t
ORIGIN
Query Match 27.2%; Score 571.4; DB 14; Length 697;
Best Local Similarity 89.0%; Pred. No. 5.1e-151;
Matches 617; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 491 GCAGCTGTTTAAAGAGTCACAGAGGATGAAGGTGCTATCACTGCAAGGCCA 550
|||||

```

```

Db 697 CACGCTGTTTATGAAAGAGTCACGAGAGGATGAAGCGCTATCACTGCAGACCCA 638
QY 551 CCACCCAGAGGGCTCTGTGAAAGTTCAGCATACCTCACTGTTCAGGAACTTCGGACA 610
Db 637 CCACCCAGAGGGCTCTGTGAAAGTTCAGCATACCTCACTGTTCAGGAACTTCGGACA 578
QY 611 AGTCTATCTGAGAGCTGATCAGCTTAAGATCACTGTGTGGCTGGACCTGCTTCGGC 670
Db 577 AGTCAGCTGGAGCTGATCAGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 518
QY 671 TCTATTAACCCCTCTTATCCGAAATGAAAGGCTCTTCTTCTGAAATGAAGAGTACT 730
Db 517 TCTGCTAACCCCTCTTATCCGAAATGAAAGGCTCTTCTGCTGAAATGAAGAGTACT 458
QY 731 ACTATCAATATTAAGGACCCAGATGAAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCC 790
Db 457 ACCCTGCAATATATCATGGACCCAGAGTCCCTGCTGATGAGCAGTCCGACGGCTGC 398
QY 791 CTATGATGCCAGCAGAGTGGAGTTCCTGCGGAGAGACTTAACTTGGCCTAAATCACTTG 850
Db 397 CTATGATGCCAGCAGAGTGGAGTTCCTGCGGAGAGACTTAACTTGGCCTAAATCACTTG 338
QY 851 GAAGAGGGCTTTTGGAAAGTGGTCAAGCATCAGCATTTGGCATTAAGAAATCACTTA 910
Db 337 GAAGAGGGCTTTTGGAAAGTGGTCAAGCATCAGCATTTGGCATTAAGAAATCACTTA 278
QY 911 CGTGGCGAGCTGTGCTGTGAAATGCTCAAGAGGGGCGCCAGCGCAGCAGTACAAAG 970
Db 277 CGTGGCGAGCTGTGCTGTGAAATGCTCAAGAGGGGCGCCAGCGCAGCAGTACAAAG 218
QY 971 CTCTGATGATGAGCTTAAATCTTGACCCCATTTGACCCCATTTGACCCCATTTGACCC 1030
Db 217 CCTGATGATGAGCTTCAAGATCTTGACCCCATTTGACCCCATTTGACCCCATTTGACCC 158
QY 1031 TCTGCGAGCGCTGACCAACCAAGAGGAGGCGCTCTGATGCTGATGCTGATGCTGATGCT 1090
Db 157 TCTGCGAGCGCTGACCAACCAAGAGGAGGCGCTCTGATGCTGATGCTGATGCTGATGCT 98
QY 1091 ATGGAATCTCTCACTACTCTCAAGAGCAAGCTGACTTATTTTCTCAACAGAGATG 1150
Db 97 ATGGAATCTCTCACTACTCTCAAGAGCAAGCTGACTTATTTTCTCAACAGAGATG 38
QY 1151 CAGCACTACATCGAGCCCTTGAAGAAAGAAAAA 1183
Db 37 CAGCTTAACTGCTGAGCCAAAAA 5

RESULT 4
AL042601
LOCUS AL042601 666 bp mRNA linear EST 29-FEB-2000
DEFINITION DKFZp341012l_r1_434 (synonym: htes3) Homo sapiens cDNA clone
AUTHORS DKFZp341012l 5', mRNA sequence.
VERSION AL042601
KEYWORDS EST.
SOURCE AL042601.1 GI:5422053
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 666)
AUTHORS Blum,H., Bauersachs,S., Mewes,H.W., Gassenkater,J. and Wiemann,S.
TITLE EST (Blum, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Blum H
MIPS
Am Kleferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.

```

This clone (DKFZp34L0121) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
1..666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp34L0121"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site:1: NotI; Site:2: SalI"
BASE COUNT 195 a 146 c 157 g 164 t 4 others
ORIGIN

Query Match 26.7%; Score 561.4; DB 9; Length 666;
Best Local Similarity 99.8%; Pred. No. 3.4e-148;
Matches 562; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1539 GAAATGATGCTCCGAAATCTATTTGACAAAATCTACAGACCAAGAGGACGCTG 1538
1 GAAATGATGCTCCGAAATCTATTTGACAAAATCTACAGACCAAGAGGACGCTG 60
1599 GTCTACGAGATGATGCTGGGAAATCTTCCTTAGGTGGTCTCCATCCAGGAG 1638
61 GCTACGAGATGATGCTGGGAAATCTTCCTTAGGTGGTCTCCATCCAGGAG 120
1659 ACAATGATGAGGACTTTTCAGTCCGCTGAGGAGGCAAGGATGAGAGCTCTGA 1718
121 ACAATGATGAGGACTTTTCAGTCCGCTGAGGAGGCAAGGATGAGAGCTCTGA 180
1719 GACCTACTCTGAAATCTATCATGATCATGCTGGACTGCTGGCAGAGACCCAAAGA 1778
181 GACCTACTCTGAAATCTATCATGATCATGCTGGACTGCTGGCAGAGACCCAAAGA 240
1779 AAGCCCAAGATTCAGAACTTTGTTGAAAATAGTGTGATTCCTCAAGCAATGTACA 1836
241 AAGCCCAAGATTCAGAACTTTGTTGAAAATAGTGTGATTCCTCAAGCAATGTACA 300
1939 ACAGATGTTAAAGACTACATCCATCAATCCATCAATCCATCAATCCATCAAT 1896
301 ACAGATGTTAAAGACTACATCCATCAATCCATCAATCCATCAATCCATCAAT 360
1899 ATACTCACTCTGCT 1958
361 ATACTCACTCTGCT 420
1959 TAATCAGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2018
421 TAATCAGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
2019 AAGAAATCAAACTTTGAGAACTTTTACCGAATCCCACTCCCACTCCCACTCCCA 2078
481 AAGAAATCAAACTTTGAGAACTTTTACCGAATCCCACTCCCACTCCCACTCCCA 540
2079 GGGCGACAGCAGCAGCTCTGTGG 2101
541 GGGCGACAGCAGCAGCTCTGTGG 563

RESULT 5
P345458
LOCUS
DEFINITION
MR3-GN0186-241100-008-c10 GN0186 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG005458
VERSION
BG005458.1 GI:12447642
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 650)

AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Borfin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jorgensen, C.V., O'Hare,
M.J., Soares, F., Brantani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE
Shotgun sequencing of the human transcriptome with cDNA expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
2002663
COMMENT:
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Frudente 109, 4 andar, 01505-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICE Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3st2-MR3-GN0186-
241100-008-c10st3-2000-11-24st4=1)
Seq primer: puc 18 forward
High quality sequence stop: 594.
Location/Qualifiers
1..650
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0186"
/dev_stage="Adult"
/note="Organ: placenta; normal; Vector: puc18; Site:1: SmaI
products derived from ORESIES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 195 a 146 c 166 g 138 t 2 others
ORIGIN
Query Match 26.6%; Score 558.2; DB 12; Length 650;
Best Local Similarity 96.6%; Pred. No. 2.8e-147;
Matches 569; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 261 GGAAGAAATCTCCAGAGAAAGAAATTCATCATCAGAGATCAGAGACCAATACCTCC 340
DB 25 GGAAGAAATCTCCAGAGAAAGAAATTCATCATCAGAGATCAGAGACCAATACCTCC 84
QY 341 TGGAAACCTCAGTATACACAGATGAGGATGAGGATGAGGATGAGGATGAGGATG 400
DB 85 GGGAAACCTCAGTATACACAGATGAGGATGAGGATGAGGATGAGGATGAGGATG 144
QY 401 CTATGGTGTCCCGAGGCTCAGTACATCTGTTTAAACCAACCAACCAACCAACCAAG 460
DB 145 CTATGGTGTCCCGAGGCTCAGTACATCTGGGTTTAAACCAACCAACCAACCAAG 204
QY 461 AGCTGGATATATATAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 520
DB 205 AGCCGGATATATATAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 264
QY 521 AGCATGAGGTGTCTATCTACTGCAAGGCAACCAACCAACCAACCAACCAAGGCTCT 580
DB 265 GCAATGAAAGGCTGCTATCAGTCAAGGCAACCAACCAACCAACCAAGGCTCT 324
QY 581 CATACCTCAGTCTCAGGACCTCGGACAGTCTAATCTGAGCTCTAATCAGCTCTAAT 640
DB 325 CATACCTCAGTCTCAGGACCTCGGACAGTCTAATCTGAGCTCTAATCAGCTCTA 384
QY 641 GCACCTGTGTGGCTCGGACTCTCTGTGGCTCTCTATTAACCTCTCTATCCGAAAAA 700
DB 365 GCACCTGTGTGGCTCGGACTCTCTGTGGCTCTCTATTAACCTCTCTATCCGAAAA 444
QY 701 AAGGTCTCTCTCTGAAATAAAGACTGACTATCTATCAATATATATGAGCCAGATGAG 760


```

Db 445 AAAGTCTCTTCGGAATTAAGACTGACTACCTATCAATATATAATGACCCAGAGAG 504
Qy 761 TTCTTTGGATGAGCAGTGTGCGGCTCCCTTATGATGCGACCAAGTGGAGTTCGCC 820
Db 505 TTCTTTGGATGAGCAGTGTGCGGCTCCCTTATGATGCGACCAAGTGGAGTTCGCC 564
Qy 821 GGAGAGAGCTTAAGTGGGCAATCACTTGGNAGAGGGGCTTTTGGAAA 869
Db 565 GGAGAGAGCTTAAGTGGGCAATCACTTGGNAGAGGGGCTTTTGGAAA 613

RESULT 6
AL575294/c 1000 bp mRNA linear EST 16-FEB-2001
LOCUS AL575294 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1061Y018 3
DEFINITION prime, mRNA sequence.
ACCESSION AL575294
VERSION AL575294.1 GI:12936322
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
source
1..1000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1061Y018"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
SE COUNT 242 a 199 c 230 g 320 t 9 others
IGIN

Query Match 22.2%; Score 466.8; DB 9; Length 1000;
Best Local Similarity 99.6%; Pred. No. 3.6e-121;
Matches 468; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TATATCAGATGTGCCAATGGGTTTCATGTTAACTTGGAAAAATGCGCAGGAAGA 60
Db 934 TATATCAGATGTGCCAATGGGTTTCATGTTAACTTGGAAAAATGCGCAGGAAGA 875
Qy 61 GAGGACCTCAACTCTCTTCAGAGTACAACTCTTATACAGACAGCTTACGATT 120
Db 874 GAGGACCTCAACTCTCTTCAGAGTACAACTCTTATACAGACAGCTTACGATT 815
Qy 121 TTACTCGGACAGTTTAAATACAGAAATGCGACACAGTATTACAGCAAAATGCCC 180
Db 814 TTACTCGGACAGTTTAAATACAGAAATGCGACACAGTATTACAGCAAAATGCCC 755
Qy 181 ATACATAGGAGCACTCCATCCTCTTAATCTTACCTATCAATGAATGTTCCCGCAAGAT 240
Db 754 ATACATAGGAGCACTCCATCCTCTTAATCTTACCTATCAATGAATGTTCCCGCAAGAT 695
Qy 241 TCAGGACACCTATGCTTCGAGAGCCAGGAATGTATATACAGGGGAGAAATCTCCAGAG 300

```

```

Db 694 TCAGGACACCTATGCTTCGAGAGCCAGGAATGTATATACAGGGGAGAAATCTCCAGAG 635
Qy 301 AAAGAAATTAACATACAGATCAGGAAGCAACATACCTCTCGGAAACCTCAGTATCAC 360
Db 634 AAAGAAATTAACATACAGATCAGGAAGCAACATACCTCTCGGAAACCTCAGTATCAC 575
Qy 361 AGATGGGCATCAGCAGTTCCACCCTTTAGACTGTGATGCTTAATGGTGTCCCGAGCTT 420
Db 574 AGATGGGCATCAGCAGTTCCACCCTTTAGACTGTGATGCTTAATGGTGTCCCGAGCTT 515
Qy 421 CAGATCACTTGGTITTAATAAACACCAACCAATATACACAGAGGCTTGGAT 470
Db 514 CAGATCACTTGGTITTAATAAACACCAACCAATATACACAGAGGCTTGGAT 465

RESULT 7
BQ770907 566 bp mRNA linear EST 26-JUL-2002
LOCUS BQ770907
DEFINITION IMAGE:6400365 5', mRNA sequence.
ACCESSION BQ770907
VERSION BQ770907.1 GI:21979381
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 566)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Issue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..566
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6400365"
/clone_lib="NIH_BMAP_F10"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (11 phage resistant)"
/notes="Organ: Brain; Vector: pYX-5; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a NotI site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-5 vector. The library tag
sequence located between the NotI site and the polyA tail
is CAGCCAGCAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
Developing Mouse Nervous System, supported by National
Institute of Mental Health (NIMH), Hsien Chin, Ph.D.,
program coordinator."
BASE COUNT 149 a 146 c 154 g 116 t 1 others
IGIN

Query Match 21.5%; Score 451.6; DB 14; Length 566;

```


/db_xref="taxon:9606"
/clone="CS0D103YL03"
/clone_lib="171_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site.1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang life technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fulllength@lifestock.com URL : <http://fulllength.lifestock.com>

Db 719 TCTCCATCGAGAGAAATGGTTAAGATCTGTGACTTCGGGTGGCCCGGACCAITAT 778
 QY 1494 TAAGAACCCCGATTATGTGAGAAAAGGAGATCTCGACTTCCTCTGAAATGGATGG 1549
 Db 779 AAGA--CCGATTATGTCAGAAAAGGAGATCCCGAATCCCTTTGAAGGTAIGG 832

Search completed: December 8, 2002, 12:49:57
 Job time : 2741.79 secs